

Ensemble Dynamics and Bred Vectors

Nusret Balci¹, Anna L. Mazzucato², Juan M. Restrepo^{3*}, George R. Sell⁴

¹ *Institute for Mathematics and its Applications*

University of Minnesota, Minneapolis MN 55455 U.S.A.

² *Department of Mathematics*

The Pennsylvania State University, University Park, PA 16802 U.S.A.

³ *Department of Mathematics and Department of Physics*

University of Arizona Tucson, AZ 85721 U.S.A.

⁴ *School of Mathematics*

University of Minnesota, Minneapolis MN 55455 U.S.A.

August 25, 2011

Abstract

We introduce the new concept of an EBV to assess the sensitivity of model outputs to changes in initial conditions for weather forecasting. The new algorithm, which we call the *Ensemble Bred Vector* or EBV, is based on collective dynamics in essential ways. As such, it keeps important geometric features which are lost in the earlier bred vector algorithm (BV). By construction, the EBV algorithm produces one or more dominant vectors, and is less prone to spurious results than the BV algorithm. It retains the attractive features of the BV with regard to being able to handle legacy codes, with minimal additional coding.

We investigate the performance of EBV, comparing it to the BV algorithm as well as the finite-time Lyapunov Vectors. With the help of a continuous-time adaptation of these algorithms, we give a theoretical justification to the observed fact that the vectors produced by BV, EBV, and the finite-time Lyapunov vectors are similar for small

*Corresponding Author: restrepo@math.arizona.edu

amplitudes. The continuum theory is established the relationship between the two algorithms and general directional derivatives.

Numerical comparisons of BV and EBV for the 3-equation Lorenz model and for a forced, dissipative partial differential equation of Cahn-Hilliard type that arises in modeling the thermohaline circulation, demonstrate that the EBV yields a size-ordered description of the perturbation field, and is more robust than the BV in the higher nonlinear regime. The EBV yields insight into the fractal structure of the Lorenz attractor, and of the inertial manifold for the Cahn-Hilliard-type partial differential equation.

Keywords: Bred vectors, Lyapunov vectors, sensitivity, dynamic stability, Cahn-Hilliard, Lorenz.

PACS 94.05sx, 92.60.Ry, 92.60.Wc, 93.65.+e, 92.70.-j, 02.50.-r

1 Introduction

Central to weather prediction is the analysis of the sensitivity of a physical or computer-coded model to initial conditions. Model sensitivity to parameters is also important in model inter-comparison. One studies such sensitivity in order to obtain a better understanding of the role played by these parameters in model outcomes.

Sensitivity and predictability are often intertwined in the context of weather prediction and have been the subject of extensive research (see Buizza et al. (1993) and references contained therein.) These are not exclusively weather-related issues and thus geophysical fluid dynamics will often mine other physical, computational and mathematical disciplines, for ideas with which to assess dynamic sensitivity. Practical sensitivity methodologies must contend with the evolution and dynamics of highly coupled, complex, high-dimensional systems, riddled with subscale parameterizations and empirical relations, which are the norm in large-scale climate and meteorology models.

A tool used in the study of sensitivity analysis is the Bred Vector (BV) algorithm. It is proposed for use in forward sensitivity of weather and climate models. While in Subsection 2.1, we present a brief survey of some of the applications of this algorithm in various sensitivity analyses, in this article, we will focus on the issue of the maximal growth of errors due to small changes in the initial conditions.

The concept of the BV algorithm we use is based on the theory first introduced in Toth and Kalnay (1993). In addition to the BV notion, we present here a new variant, which we call the *Ensemble Bred Vector* (EBV) algorithm. The definitions of both the BV and EBV algorithms are presented in Section 2.

In the BV algorithm, one follows an initial condition of the time-discrete nonlinear system, along with cloud, which describes a family of nearby solutions. (Since this algorithm is used to sample the error space, an ensemble of initial perturbations is bred simultaneously.) The perturbations at the initial time are fixed with a common small amplitude ϵ . After each cycle, the outcome of the perturbations is rescaled to the same amplitude ϵ . For the BV algorithm, the rescaling of each perturbation is independent of the others, and there is no mechanism to use the rescaling to compare the dynamics of nearby perturbations.

The new variation that we propose here, the EBV algorithm, differs from the BV algorithm, in the rescaling rule. In particular, for the EBV those perturbations that are not the same size as the largest perturbation, play a reduced role after the rescaling. Thus the rescaling used in the EBV algorithm serves us better in separating various levels of the dominant dynamics. In short, the EBV algorithm offers better insight into the relative behavior of nearby trajectories. Therefore, even when the initial perturbations of the two algorithms are based on the same cloud, the EBV algorithm is linked to the ensemble dynamics of the underlying non-linear model, hence its name. We will make use of both the BV and the EBV algorithms. In fact, one of our major goals is to present an in-depth comparison of the two algorithms, as they are used, or can be used, in describing the underlying dynamics of the model.

We will show that in some metrics, the BV and EBV algorithms are comparable, with the EBV being more accurate and faster, see Table 1 and Figures 6, 7, 8, for example. For other issues, especially those involving the longtime dynamics within the global attractor, the BV algorithm has a shortcoming, which limits its use (see Section 4). The EBV, instead, leads to useful and interesting insight into the dynamics of the model, as is shown in the three Figures 4, 1, and 2.

This brings up the question: For a given model of sensitivity with respect to initial conditions, how does one determine the direction vector that results in the maximal increase in the error due to a small perturbation in the given direction initially? This is where the Lyapunov vectors enter the scene. What one needs is a **red vector**, which is the Lyapunov vector x_0 with $\|x_0\| = 1$ and with the property that the corresponding strong Lyapunov exponent λ_1 is the maximal Lyapunov exponent for the model. (See Subsection 2.5 for the definition and more details. One should note that the Lyapunov exponents for the model require integration over $0 \leq t < \infty$, or over the real line \mathbb{R} .)

We will use either the EBV or the BV algorithm to approximate the solutions of the tangent linear equation. In Section 3 we show that either algorithm is a good approximation. For these algorithms one can integrate only over a finite interval $0 \leq t \leq T$, where $0 < T < \infty$. However, it

is only under exceptional circumstances, *e.g.*, time-periodic or autonomous problems, that a finite-time integration will approximate well an infinite time average. For any hope for success in using a finite-time integration, we require that the model problem satisfy two properties:

- (1) There is an attractor \mathfrak{A} for the solutions,
- (2) The initial condition is chosen to be on, or near, \mathfrak{A} .

What can one expect with such a finite-time approximation, when the initial conditions are *near* the attractor \mathfrak{A} ? In terms of the calculated time, one expects first to be in a transient state. Then after a while, one hopes to get some meaningful information about the longtime dynamics of the model. We include in this manuscript several studies of such approximations.

It is very important to note that, in order to better understand the maximal growth of errors due to small changes in the initial conditions, one needs to exploit the dynamical information contained in the attractor of the model. In particular, one needs to complete two steps:

- Step 1: One needs to locate the red vector x_0 and the associate Lyapunov exponent λ_1 . This then determines the red vector solution $R(x_0, t)$, for $-\infty < t < \infty$, for the model equation.
- Step 2: One must find a good approximation of the time evolution $R(x_0, t)$ on an appropriate finite-time interval, $0 \leq t \leq T < \infty$.

Once the red vector x_0 is known, then ϵx_0 is the initial condition for the bred vector sequence, see Subsection 2.2. Due to the results derived here in Section 3, either $EBV(t)$, or $BV(t)$, is a good finite-time approximation of $R(x_0, t)$, for $0 \leq t \leq T < \infty$. This takes care of Step 2. Consequently, the problem boils down to a search for the red vector, which is addressed below in Subsection 2.5. Fortunately for us, there is wealth of related mathematical information in the 1987 manuscript of R. A. Johnson, K. J. Palmer, and G. R. Sell. see Johnson et al. (1987). (We will refer to this paper as the “JPS87” in the sequel.) As we shall show, by using the JPS87, we are able to describe the mathematical process of finding the red vector. By using this citation, with the theory of the $EBV(t)$ algorithm, this leads to a good solution for our sensitivity problem.

It was observed by Toth & Kalnay *op. cit.* (see also Toth and Kalnay (1997)) in several experimental runs that BVs resemble the *leading* finite-time Lyapunov vectors. In order to make a more quantitative comparison between Lyapunov vectors and BVs, in Section 3 we study the Continuum Limits (as the basic time-step size for rescaling goes to 0) of the BV and the EBV algorithms and show direct connections between these limits and

specific solutions of the continuous-time tangent linear equations. Section 3 also contains a further discussion of some of the desirable and interesting features of the new EBV. For instance, there is a natural ordering of the ensemble members of an EBV, and as we will show, it is possible to observe perturbations with smaller sizes than the dominant one, but with very strong growth, see the spear-like behavior in Figure 4.

We consider two models to exemplify the features of the new EBV. The first is the familiar Lorenz63 model introduced by Lorenz (1963). It has a well-known global attractor. The second is a nonlinear forced and dissipative partial differential equation of the Cahn-Hilliard type. This equation is a variation of a model proposed by Cessi and Young (1992) of the oceanic thermohaline circulation. We will denote the equation associated with this model as the CY92 in this study. (In fact, we impose periodic boundary conditions instead of the more physical zero-flux, zero-stress conditions at the poles.)

It turns out, it is a good example of the typical climate-related model dynamics. However, the CY92 is special, since it has an inertial manifold. Consequently, the longtime dynamics of this partial differential equation is completely contained in the attractor of a finite dimensional ordinary differential equation. In applications, the BV algorithm and its variants are in fact discrete-time algorithms based on finite-dimensional approximations of weather models, obtained either by mode projection as in the Lorenz (1963) system, or by spatial discretization of partial differential equations (PDEs), as in the Cessi-Young (CY92) model, which will be described below. As we will show the EBV yields insights into the structure of the attractor of the Lorenz63, and of the inertial manifold to the CY92.

Our numerical examples also highlight that the BV algorithm is sensitive to the amplitude and frequency content of the initial perturbation. In contrast, the outcome of the EBV algorithm shows a clear hierarchy among its members, and the first few members already generate an unambiguous characterization of the perturbation field at both large and small amplitudes. In the nonlinear regime, however, the EBV will be shown in Section 4 to be less likely to produce spurious results than the BV. In Section 5 we will address implementation issues of the EBV.

2 BV and EBV Algorithms; Finite-Time Lyapunov Vectors

In this section we present the definitions and methodology for computing the two Bred Vector algorithms, the BV and the EBV. We also review the basic theory of Lyapunov Vectors, Lyapunov exponents, and their finite-time

counterparts. We will compare these different tools below. However, before doing this, we include here a brief survey of some of the applications and theoretical issues that have been noted in the implementation of the BV algorithm.

2.1 Brief Survey of Applications of the BV Algorithm

For the BV algorithm, we use the one originally proposed by Toth and Kalnay (1993). BV is purely algorithmic. It is “equation-free” and thus with additional minimal computer coding it can handle legacy code representing even extremely complex models. Most alternatives for obtaining estimates of forward sensitivity will involve non-trivial additional coding. For example, in order to obtain the finite-time Lyapunov vectors and exponents, one needs to derive and make use of the tangent linear model. (From the beginning, it was realized that there existed some close connections between the BV algorithm and the tangent linear equations. As we will show in Section 3, there is a rigorous mathematical foundation for these connections.) Singular value decomposition methods, which can offer complementary information to Lyapunov-vector inspired methods, also require a tangent linear model. Deremble et al. (2009) use this approach to study regime predictability in some reduced weather models. Other examples are: Buizza et al. (1993) and Palmer et al. (1998). In another direction, Wolfe and Samelson (2007) propose the use of the MET and the finite-time singular vectors to approximate the Lyapunov vectors.

The BV algorithm is a finite-time, forward sensitivity methodology which, in addition to being useful in characterizing model sensitivity to initial conditions, has been proposed as a means to produce a reduced-rank representation of the background error in data assimilation and forecast error-covariance approximations (see Corazza et al. (2003), for example).

Several articles in the literature have addressed applications of the BV algorithm in weather modeling. See, in particular Toth and Kalnay (1993), Toth and Kalnay (1997), Kalnay (2003), and reference therein. For a comparison of the BV algorithm and other methods, such as Monte-Carlo perturbed observations, we refer for example to Cheung (2001); Gneiting and Raftery (2005); Hansen and Smith (2000); Wei and Toth (2003). For an application of the BV algorithm to ensemble Kalman filters see Wang and Bishop (2003). Primo et al. (2008) and Hallerberg et al. (2010) have treated applications based on variations in the algorithm, where for example, the rescaling is done by using a geometric mean.

As already discussed, several alternatives to the BV algorithm have been proposed and employed in the literature. BVs have been viewed as non-linear analogs of finite-time Lyapunov vectors. Similarly nonlinear analogs of sin-

gular vectors have been proposed, for instance conditional nonlinear optimal perturbations proposed by Mu and Jiang (2008) and non-linear singular vectors proposed by Rivière et al. (2008), although they entail a computationally expensive optimization.

Both the BV and EBV algorithms arise in the time-discretization of a continuous-time dynamical system. We consider the following initial value problem

$$\begin{aligned}\frac{dy}{dt} &= G(y), \quad t > 0, \\ y(0) &= y_0,\end{aligned}\tag{1}$$

where t represent time and $G = G(y)$ is a map that has at least a bounded gradient. Since our main applications involve autonomous differential equations, we assume that G does not explicitly depend on time. The basic theory we present here has a routine extension to non-autonomous problems.

The solution vector $y = y(t)$ can live in a finite- or infinite-dimensional normed linear space. In the former case, (1) is an (autonomous) system of ordinary differential equations, while in the latter case, (1) is an (autonomous) system of partial differential equations, modeling a time dependent, spatially extended system. For systems of partial differential equations, we assume that either periodic boundary conditions or non-flux boundary conditions are prescribed. See for example, Sell and You (2002). The use of other boundary conditions may lead to a related theory, but we do not address the issue here.

If the system contains evolution partial differential equations, the system, along with the boundary conditions, are discretized in space or projected onto a finite-dimensional space compatible with the boundary conditions. Consequently, we usually assume that (1) is a system of ordinary differential equations of dimension K , which may be large. Since most large-scale weather and climate circulation models presently use explicit-in-time integrators, we will focus on numerical models of this type.

2.2 Bred Vector Algorithms: BV and EBV

We let $y = y(t)$ denote a given (continuous time) solution of (1). We then turn to an approximate solution $Y_n = Y(t_n)$, which is defined on the time grid: t_n , for $n = 0, 1, 2, \dots$, where $t_{n+1} = t_n + \delta t_n$, for $n = 0, 1, 2, \dots$. We set $t_0 = 0$. We assume that $\delta t_n = \delta t$ is positive and small, and that it does not depend on n , for $n \geq 0$.

For the autonomous case, the initial value problem, which is approximated using an explicit numerical integration scheme, leads to consideration of the

difference equation

$$\begin{aligned} Y_{n+1} &= M(Y_n, \delta t), \quad n = 0, 1, 2, \dots, \\ Y_0 &= Y(0). \end{aligned} \tag{2}$$

where $Y(t_n)$ is a solution of the discrete problem (2), and it may be viewed as an approximation of the continuous-time solution $y(t)$, at $t = t_n$. Likewise, the initial condition Y_0 is an approximation of the initial condition y_0 . The points Y_n are in the K -dimensional Euclidean space \mathbb{R}^K and $M = M(Y, \delta t)$ is the discrete-time solution operator on \mathbb{R}^K generated by the ordinary differential equation (1). As noted above, we assume that the discrete-time problem (2) has an attractor \mathfrak{A} , and that Y_0 is on, or near \mathfrak{A} .

At this point it is convenient to introduce the related concepts of a *Cloud* (at $t = 0$) and a *Family of Initial Perturbations*. A Cloud (at $t = 0$) is a family of tangent vectors $\delta\mathcal{Y}_0(\iota)$ to Y_0 in \mathbb{R}^K that depends on ι , where $\iota \in \mathbb{I}$ and \mathbb{I} is a finite index set. The main requirement we impose is that

$$\|\delta\mathcal{Y}_0(\iota)\| = \epsilon, \quad \text{for all } \iota \in \mathbb{I},$$

where ϵ is small, positive, and fixed. The collection of all terms $(Y_0, \delta\mathcal{Y}_0(\iota))$ in $\{Y_0\} \times \mathbb{R}^K$, for $\iota \in \mathbb{I}$, is called a Family of Initial Perturbations. Notice that this collection lies on a sphere of radius ϵ in $\{Y_0\} \times \mathbb{R}^K$ with center at $(Y_0, 0)$. Both the Cloud and the perturbations evolve in time, via the BV or the EBV algorithms, which are defined below. As we now note, the definitions of the two algorithms differ only in the rescaling rule.

We begin by recalling the BV algorithm as given by Toth&Kalnay. For $n \geq 0$, we assume that the base point Y_n and the perturbation vector $\delta Y_n(\iota)$ are known. For the $(n+1)^{\text{st}}$ step we use:

1. Y_{n+1} denotes the $(n+1)^{\text{st}}$ base point, and it is determined by (2);
2. $\delta\mathcal{Y}_{n+1}(\iota)$, the $(n+1)^{\text{st}}$ perturbation vector, is given by

$$\delta Y_{n+1}(\iota) = M(Y_n + \delta Y_n(\iota), \delta t) - M(Y_n, \delta t), \tag{3}$$

$$\delta\mathcal{Y}_{n+1}(\iota) = R_{n+1} \delta Y_{n+1}(\iota), \tag{4}$$

where R_{n+1} is a rescaling rule.

The time evolution of the Cloud is $\text{BV}(t_n) = \delta\mathcal{Y}_n(\iota)$, where $\iota \in \mathbb{I}$ and $n \geq 0$.

One rescaling rule, proposed by Toth and Kalnay, *op. cit.*, consists of rescaling the n^{th} perturbation vector to the previous one by

$$\|\delta\mathcal{Y}_{n+1}(\iota)\| = R_{n+1} \|\delta Y_{n+1}(\iota)\| = \|\delta\mathcal{Y}_0(\iota)\| = \epsilon, \quad \text{for } \iota \in \mathbb{I}, n \geq 0. \tag{5}$$

Equivalently,

$$R_{n+1}(Y_0) := \frac{\|\delta\mathcal{Y}_n(\iota)\|}{\|\delta Y_{n+1}(\iota)\|} = \frac{\|\delta\mathcal{Y}_0(\iota)\|}{\|\delta Y_{n+1}(\iota)\|}, \tag{6}$$

so that $R_{n+1} = R_{n+1}(Y_0)$ depends on the initial base point Y_0 , as well as the initial perturbation vector $\delta\mathcal{Y}_0(\iota)$. An alternate rescaling rule consists in rescaling periodically, at t_{mk} , where k is an integer $k \geq 2$, and $m = 1, 2, \dots$. In this case, one uses the rule (6) when $n = mk$, and

$$R_{n+1}(Y_0) = 1, \quad \text{for } mk < n < (m+1)k. \quad (7)$$

We do not use this alternate rule in this paper. (For a discussion of rescaling time and regime predictability in some reduced models, see Deremble et al. (2009) and references therein).

For the *Ensemble Bred Vector* algorithm, instead of using R_{n+1} as in equation (3), we use a uniform scaling \mathbf{R}_{n+1}^{\min} , which is the same for all $\iota \in \mathbb{I}$. In particular, we replace equation (3) with

$$\begin{aligned} \delta Y_{n+1}(\iota) &= M(Y_n + \delta\mathcal{Y}_n(\iota), \delta t) - M(Y_n, \delta t) \\ \delta\mathcal{Y}_{n+1}(\iota) &= \mathbf{R}_{n+1}^{\min} \delta Y_{n+1}(\iota), \end{aligned} \quad (8)$$

for all $\iota \in \mathbb{I}$, where

$$\mathbf{R}_{n+1}^{\min} = \epsilon \left[\max_{\iota \in \mathbb{I}} (\|\delta Y_{n+1}(\iota)\|) \right]^{-1}. \quad (9)$$

Similarly, when (8) and (9) hold, we use $\text{EBV}(t_n) = \delta\mathcal{Y}_n(\iota)$, for $\iota \in \mathbb{I}$ and $n \geq 0$, to denote the time evolution of the Cloud. (Alternatively, a periodic rescaling rule utilizing (7) above can be employed.) The time step used to compute the base trajectory and the time intervals between normalizations need not be the same. This is the case for both the BV algorithm as well as EBV.

The crucial difference between the BV and EBV algorithms is that, even when the BV is run concurrently over an ensemble of initial data, the outcome of the algorithm for each given datum does not depend on the other members of the ensemble. In contrast, the evolution of the ensemble members is interdependent in the EBV. Nevertheless, by construction, the EBV should exhibit similar behavior to the BV at small amplitudes. Indeed, one of the design principles for the EBV algorithm is to reinforce this aspect by providing it with a built-in acceleration mechanism.

Both the BV algorithm and EBV outcomes, on the other hand depend on the choice of vector norm used to define the rescaling rule in either (6) or (9). While all norms are equivalent in the (finite) K -dimensional space where we seek solutions, in practice the constants appearing in the equivalence between different finite-dimensional norms generally strongly depend on the dimension K and eventually blow up as k becomes infinite. Hence, the choice of norm used can have an impact on the implementability and performance of these algorithms. In Rivière et al. (2008), it was suggested that the outcome

of the BV algorithm may depend strongly on the choice of norm. This is actually a consequence of the non-negligible nonlinear effects in the system. In Section 5, we will elaborate further on the issue of norm dependence.

In addition to the “renormalization time-step” δt , there is another time step, the “integration time-step”, which we will denote by Δt . For example, one encounters the new time step when moving from the continuous-time problem equation (1) to the discrete-time problem equation (2). For the most part, we will treat Δt and δt as being equal in the calculations described in this article. However, we always require that $\Delta t \leq \delta t$.

2.3 Lyapunov Vectors

In Section 4, we will be applying the BV and EBV algorithms to two models consisting of systems of (nonlinear) autonomous ODEs of the form (1). (The second model arises from the discretization of a PDE.) In each model, there is a compact, global attractor \mathfrak{A} , which is a subset of \mathbb{R}^K , and is invariant for the time evolution of the system. (We refer the reader to Chapter 2 in Sell and You (2002) for more information on attractors and global attractors.) We will let θ denote a typical point in the attractor \mathfrak{A} , and we will let $\theta \cdot t = y(t)$ denote the unique solution of (1) that satisfies $y(0) = \theta$. Since \mathfrak{A} is invariant, one has $\theta \cdot t \in \mathfrak{A}$, for all $t \in \mathbb{R}$.

In order to study the sensitivity with respect to initial conditions on the attractor, the Tangent Linear Model is used, which is defined as

$$\partial_t x = A(\theta \cdot t) x, \quad (10a)$$

$$x(0) = x_0 \in \mathbb{R}^K, \quad (10b)$$

where $A(y) = DG(y)$ is the Jacobian matrix of G . Hence $A = A(\theta \cdot t)$ is the linearization of (1) along the solution $\theta \cdot t$. We observe that, even if G does not explicitly depend on time, (10a) is generally non-autonomous, since $\theta \cdot t$ changes with time.

We let $U(\theta, t)$, denote the solution operator of (10a), which takes the initial data to the solution at time t , so that $U(\theta, t)x_0$ is the solution of the initial value problem for (10). Such an operator is well defined by the uniqueness of solutions to the problem (10). Uniqueness of solutions also readily implies the cocycle identity:

$$U(\theta, \tau + t) = U(\theta \cdot \tau, t) U(\theta, \tau), \quad \text{for all } \theta \in \mathfrak{A} \text{ and all } \tau, t \in \mathbb{R}. \quad (11)$$

Next we consider a family of mappings $\Pi = \Pi(t)$, which are defined for $t \in \mathbb{R}$ by the relation

$$\Pi(t)(\theta, x_0) := (\theta \cdot t, U(\theta, t)x_0), \quad \text{for } (\theta, x_0) \in \mathfrak{A} \times \mathbb{R}^K. \quad (12)$$

We note that $\Pi(t)$ maps $\mathfrak{A} \times \mathbb{R}^K$ into itself, for each $t \in \mathbb{R}$; it is jointly continuous in (t, θ, x_0) ; it satisfies $\Pi(0)(\theta, x_0) = (\theta, x_0)$, (i.e., $\Pi(0) = I$, the identity operator; as well as the evolution property:

$$\Pi(\tau + t) = \Pi(t) \Pi(\tau), \quad \text{for all } \tau, t \in \mathbb{R}. \quad (13)$$

By using the discrete-time dynamics, where t and τ are restricted to satisfy $t = n \cdot \delta t$ and $\tau = m \cdot \delta t$, the notation and the theory of dynamical systems extends readily to the discrete-time problems of interest herein. Note that the θ -component of Π does *not* depend on the x_0 -component. Thus Π is called a *skew product* flow. Since Π is linear in x_0 , it is sometimes called a *linear* skew product flow. In summary, the Tangent Linear Equation over the attractor \mathfrak{A} generates a linear skew product flow. (For more information on the theory of skew product flows in the context of non-autonomous dynamics, see the multiple works of Sacker and Sell, for example: Sacker and Sell (1977), Sacker and Sell (1978) and Sacker and Sell (1980).) The dynamics of Π are crucial for understanding the sensitivity and predictability of the underlying model.

In his opus magnum, which was published in Ukraine in 1892, Lyapunov presented his theory of stability for finite-dimensional ordinary differential equations. This work includes his study of the non-autonomous linear problem (10), see Lyapunov (1992) (yes, 100 years later.) One of Lyapunov's goals was to develop an analogue of the well-known eigenvalue-eigenvector theory, for the solutions of the autonomous problem, to the study of solutions of general non-autonomous equation (10).

The approach developed by Lyapunov begins with the 4 Lyapunov Relations of exponential growth:

$$\limsup_{t \rightarrow \pm\infty} \frac{1}{t} \log(\|U(\theta, t)x_0\|) \quad \text{and} \quad \liminf_{t \rightarrow \pm\infty} \frac{1}{t} \log(\|U(\theta, t)x_0\|),$$

where $x_0 \neq 0$. Lyapunov was interested in, as are we, the case where these four limits are equal, and

$$\lambda(\theta, x_0) \stackrel{\text{def}}{=} \lim_{t \rightarrow -\infty} \frac{1}{t} \log(\|U(\theta, t)x_0\|) = \lim_{t \rightarrow \infty} \frac{1}{t} \log(\|U(\theta, t)x_0\|), \quad (14)$$

where $x_0 \neq 0$. The linearity of $U(\theta, t)x_0$ implies that $s\lambda(\theta, x_0) = \lambda(\theta, sx_0)$, for $s \neq 0$. Hence one can assume, as we do, that x_0 is a unit vector, i.e., $\|x_0\| = 1$. When (14) holds, then $\lambda(\theta, x_0)$ is a *strong Lyapunov exponent*, and the unit vector x_0 is an (associate) *Lyapunov vector*. The *Lyapunov spectrum*, $\text{LY}\Sigma$, is the collection of all such $\lambda(\theta, x_0)$, with $\theta \in \mathfrak{A}$ and $\|x_0\| = 1$. For example, if $A(\theta \cdot t) = A_0$ is an autonomous matrix, then the Lyapunov spectrum consists of all real numbers λ that satisfy $\lambda = \text{Re } \nu$, where ν is an eigenvalue of A_0 .

A Lyapunov vector is not an isolated vector, rather it spawns a line of Lyapunov vectors (through the origin) in \mathbb{R}^K . That is to say, a Lyapunov vector is a *point* in \mathcal{P}^{K-1} , the $(K-1)$ -dimensional projective space. For a given vector $v \neq 0$ in \mathbb{R}^K , we will use $[v] \in \mathcal{P}^{K-1}$ to denote the unique line in \mathbb{R}^K that contains v . (Note that $[v] = [-v]$.) Conversely, when one maps a line $[u]$ in \mathcal{P}^{K-1} to a vector $u \in \mathbb{R}^K$, we require that the pre-image u lie on the line $[u]$ and that $\|u\| = 1$. One should note that \mathcal{P}^{K-1} is a metric space. The *projective metric* $d_{\text{pr}}(u_1, u_2)$ is defined for nonzero vectors u^1 and u^2 in \mathbb{R}^K by

$$d_{\text{pr}}(u^1, u^2) = \min_{s_1, s_2} \|s_1 u^1 \pm s_2 u^2\|, \quad (15)$$

where s_1 and s_2 are real numbers that satisfy $\|s_1 u^1\| = \|s_2 u^2\| = 1$. This metric is used for measuring the distance between the lines $[u^1]$ and $[u^2]$ in \mathcal{P}^{K-1} .

Since the solution operator $U(\theta, t)$ of the linear problem (10) maps lines in \mathbb{R}^K onto lines, one can use this operator to define a related *projective flow* $\Sigma(\theta, t)$ on \mathcal{P}^{K-1} by means of the relation

$$\Sigma(\theta, t)[u] = [U(\theta, t)u], \quad \text{for } [u] \in \mathcal{P}^{K-1}.$$

Using this, one obtains an equivalent flow $[\Pi]$ on $\mathcal{P}^{K-1} \times \mathfrak{A}$, where

$$[\Pi](t)(\theta, [x_0]) = (\theta \cdot t, \Sigma(\theta, t)[x_0]), \quad \text{for } (\theta, [x_0]) \in \mathfrak{A} \times \mathcal{P}^{K-1}, \quad (16)$$

compare with (12). One obtains additional information about the dynamics on the projective flow $[\Pi(t)]$, by using the Lyapunov vectors, as is noted below.

2.4 The Finite-Time Lyapunov Vectors

Next we turn our attention to the question of finding good finite-time approximations of these Lyapunov vectors.

We begin by constructing a piecewise autonomous approximation of the Tangent Linear Equation for (1). To this end, we replace (10) by

$$\partial_t Y(t) = A_n Y(t), \quad \text{for } t_n < t \leq t_{n+1}, \quad (17a)$$

$$Y(0) = Y_0, \quad (17b)$$

$$t_n = n \delta t, \quad n = 0, 1, 2, \dots, N-1. \quad (17c)$$

where $A_n = A(y(t_n)) := \frac{\partial G}{\partial y}|_{y=y(t_n)}$ and $y(t_n)$ is the value of an exact solution of (1) at $t = t_n$. The solution of this system at the grid points t_n is, explicitly and recursively, given by

$$Y_{n+1} = e^{\delta t A_n} Y_n, \quad \text{for } n = 0, 1, 2, \dots, N-1. \quad (18)$$

Consequently, Y_N , the solution at time $T = N \delta t$, is given by

$$Y(T) = Y_N = Z(T) Y_0, \quad (19)$$

where

$$Z(T) := e^{\delta t A_{N-1}} e^{\delta t A_{N-2}} \dots e^{\delta t A_1} e^{\delta t A_0}.$$

We now define the approximation of the Lyapunov Vector associated with the largest Lyapunov Exponent λ_1 , at the time $t = T$ - *the finite-time Lyapunov vector*, as the direction of steepest ascent for the matrix $Z(T)$. For example, if one had used an explicit Euler scheme, then $W(T)$ would be an Eulerian approximation of $Z(T)$, where

$$W(T) = (I + \delta t A_{N-1}) \dots (I + \delta t A_1) \dots (I + \delta t A_0).$$

The finite-time Lyapunov Vector is the singular vector corresponding to the largest singular value of $Z(T)$.

In practice (see e.g. Section 4.2 for the case of the CY92 model), the finite-time LV will be computed by directly solving a discrete approximation of the LTM and rescaling the output (the rescaling can be done at arbitrary intervals of time, since the problem is linear.)

2.5 The Search for the Red Vector

One of the main contributions found in the JPS87 manuscript is an indepth study of the interactions between two major theories of the longtime dynamics of nonautonomous, linear differential systems. Dynamics of nonautonomous, linear differential equations:

- Exponential Dichotomies (and Continuous Foliations) and
- the MET (Multiplicative Ergodic Theorem) and Ergodic Measures.

We view the JPS87 manuscript as a toolkit to be used in the analysis of the dynamics of related linear systems: It is this united theory, as we now show, that forms the mathematical foundations of the theory and applications of bred vectors. We begin with the first aspect: Exponential Dichotomies and Continuous Foliations.

Consider the family of shifted semiflows

$$U_\lambda(\theta, t) \stackrel{\text{def}}{=} e^{-\lambda t} U(\theta, t), \quad \text{for } \lambda \in \mathbb{R},$$

and the associate skew-product flows

$$\Pi_\lambda(t)(\theta, x_0) \stackrel{\text{def}}{=} (\theta \cdot t, U_\lambda(\theta, t) x_0).$$

Let \mathcal{M} be a compact invariant set in \mathfrak{A} . As noted in Sacker and Sell (1978), the skew-product flow Π_λ is said to have an *exponential dichotomy* over \mathcal{M} , provided that there exist projectors \mathbb{P}_λ and \mathbb{Q}_λ and constants $K_0 \geq 1$ and $\alpha > 0$, such that $P_\lambda(\theta) + Q_\lambda(\theta) = I$, and such that, for all $\theta \in \mathcal{M}$ and $u \in \mathbb{R}^K$, one has

$$\begin{aligned} \|U_\lambda(\theta, t) Q_\lambda(\theta) u\| &\leq K_0 e^{-\alpha t} \|u\|, & t \geq 0 \\ \|U_\lambda(\theta, t) P_\lambda(\theta) u\| &\leq K_0 e^{\alpha t} \|u\|, & t \leq 0. \end{aligned}$$

When there is an exponential dichotomy and (2.5) is valid, then the *stable* and *unstable* linear spaces, \mathcal{S}_λ and \mathcal{U}_λ - which are respectively the ranges of the projectors Q_λ and P_λ - satisfy important dynamical properties that describe the exponential growth rate of selected solutions. For example, the system (2.5) is equivalent to:

$$\begin{aligned} \|U_\lambda(\theta, t) u\| &\leq K_0 e^{-\alpha t} \|u\|, & \text{for } t \geq 0, u \in \mathcal{R}(Q_\lambda(\theta)), \\ \|U_\lambda(\theta, t) u\| &\leq K_0 e^{\alpha t} \|u\|, & \text{for } t \leq 0, u \in \mathcal{R}(P_\lambda(\theta)). \end{aligned} \quad (20)$$

Let $\text{SS}\Sigma$ denote the SS Spectrum (aka Sacker-Sell Spectrum) for Π , which is defined as the collection of all $\lambda \in \mathbb{R}$ such that Π_λ does not have an exponential dichotomy over \mathcal{M} .

The Spectral Theorem in Sacker and Sell (1978) describes the continuous foliation, and other properties of the flow Π over \mathcal{M} . More precisely, there is an integer ℓ , where $1 \leq \ell \leq K$, such that $\text{SS}\Sigma$ is the union of ℓ closed, bounded intervals, that is,

$$\text{SS}\Sigma = \bigcup_{i=0}^{\ell-1} [a_{\ell-i}, b_{\ell-i}], \quad \text{with } a_{\ell-i} \leq b_{\ell-i} < a_{\ell-i-1}.$$

Also there is a continuous foliation

$$\mathbb{R}^K = \bigoplus_{i=0}^{\ell-1} V_{\ell-i}(\theta), \quad \theta \in \mathcal{M},$$

where $\{V_\ell(\theta), \dots, V_1(\theta)\}$ is a linearly independent, continuous family of subspaces of \mathbb{R}^K with $\sum_{i=0}^{\ell-1} \dim V_{\ell-i}(\theta) = K$, for $\theta \in \mathcal{M}$. As is shown below, the right-most interval $[a_1, b_1]$ plays a special role in the study of bred vectors.

For $1 \leq k \leq \ell$, we let $[a_k, b_k]$ denote the k^{th} spectral interval and let $V_k(\theta)$ denote the corresponding subspace given by the continuous foliation. We next explore the important connections between the exponential growth rates of solutions with initial conditions in $V_k(\theta)$ and the k^{th} interval $[a_k, b_k]$. Among other things, we will encounter the Monotonicity Property and the Strictly Monotone Property.

Let ν and λ be real numbers that satisfy $\nu < \lambda$, and both $U_\nu(\theta, t)$ and $U_\lambda(\theta, t)$ have exponential dichotomies over \mathcal{M} . Then the Monotonicity Property holds:

$$\mathcal{S}_\nu \subseteq \mathcal{S}_\lambda, \quad \text{while} \quad \mathcal{U}_\lambda \subseteq \mathcal{U}_\nu. \quad (21)$$

The Strictly Monotone Property is a consequence of the observation that the following three statements are equivalent:

- One has $\mathcal{S}_\nu \not\subseteq \mathcal{S}_\lambda$.
- One has $\mathcal{U}_\lambda \not\subseteq \mathcal{U}_\nu$.
- There is a spectral interval $[a_k, b_k]$ in the interval (ν, λ) .

This brings us to a basic property. Let $b_{\ell+1}$ and a_0 satisfy: $b_{\ell+1} < a_\ell$ and $b_1 < a_0$. Next we fix ν and λ so that

$$b_{k+1} < \nu < a_k \leq b_k < \lambda < a_{k-1}.$$

Then neither ν nor λ lie in the Sacker-Sell spectrum $SS\Sigma$. Furthermore, the interval (ν, λ) contains the spectral interval $[a_k, b_k]$. By the Strictly Monotone Property, the space \mathcal{S}_λ is larger than \mathcal{S}_ν , while \mathcal{U}_ν is larger than \mathcal{U}_λ . Moreover, as is shown in Sacker and Sell (1978), one has:

$$V_k(\theta) = \mathcal{U}_\nu(\theta) \cap \mathcal{S}_\lambda(\theta), \quad \text{for all } \theta \in \mathcal{M}. \quad (22)$$

It is a consequence of the relations (22) and (20) that if the initial condition (θ, x_0) satisfies $x_0 \in V_k(\theta)$, then the solution $U(\theta, t)x_0$ satisfies

$$\begin{aligned} \|U(\theta, t)x_0\| &\leq K_0 e^{(\lambda-\alpha)t} \|x_0\|, & \text{for } t \geq 0, \\ \|U(\theta, t)x_0\| &\leq K_0 e^{(\nu-\alpha)t} \|x_0\|, & \text{for } t \leq 0. \end{aligned}$$

(Note that K_0 depends on the choice of ν and λ .)

It should be noted that all the terms used above, including $\mathbb{P}_\lambda(\theta)$ and $\mathbb{Q}_\lambda(\theta)$, vary continuously in θ . Furthermore, the exponential dichotomy is robust, in the sense that it varies continuously under small perturbations. Small changes in the model result in a related exponential dichotomy with small changes in \mathbb{P}_λ , \mathbb{Q}_λ , K_0 , and α , see Pliss and Sell (1999).

Moreover, it is shown in Sacker and Sell (1976) that if $\lambda \in SS\Sigma$, then there is a $(\theta, x_0) \in \mathcal{M} \times \mathbb{R}^K$, such that $\|x_0\| \neq 0$ and the solution $U_\lambda(\theta, t)x_0$ satisfies:

$$\sup_{t \in \mathbb{R}} \|U_\lambda(\theta, t)x_0\| < \infty. \quad (23)$$

What will become apparent shortly is that the red vector must be in the space $V_1(\theta)$. Furthermore, since $\lambda = b_1$ is in $SS\Sigma$, the pair (θ, x_0) , that arises in

(23) for this choice of λ , is a candidate for the “red vector” designation. A red vector must satisfy (23), but the converse need not be true. More on this later.

As noted above, the second tool to be used in the theory of Lyapunov exponents/vectors is the Multiplicative Ergodic Theorem (MET) and the ergodic measures on \mathcal{M} . One finds in JPS87 a study of the links between the $\text{LY}\Sigma$ and the $\text{SS}\Sigma$. While the $\text{SS}\Sigma$ leads to a continuous foliation, as noted above, the MET leads to a “measurable” refinement of this continuous splitting, as is noted in Remark 4.2.9 on pages 177-179 in Arnold (1998). The latter reference is noteworthy because it contains various extensions of the MET to problems not originally envisioned in the pioneering works of an earlier generation.

The Multiplicative Ergodic Theorem: Let \mathcal{M} be a (non-empty) compact, invariant set on the attractor \mathfrak{A} , and let μ be an ergodic measure on \mathcal{M} with $\mu(\mathcal{M}) = 1$. Then there is an invariant set \mathcal{M}_μ in \mathcal{M} , with $\mu(\mathcal{M}_\mu) = 1$, and there is a k , with $\ell \leq k \leq K$, such that the following hold:

1. There is a measurable foliation

$$\mathbb{R}^K = \bigoplus_{j=0}^{j=k-1} W_{k-j}(\theta), \quad \text{for } \theta \in \mathcal{M}_\mu,$$

where $\{W_{k-j}(\theta), \dots, W_1(\theta)\}$ is a linearly independent, measurable family of subspaces of \mathbb{R}^K with $\dim W_j(\theta) = m_j \geq 1$, for $1 \leq j \leq k$, $(m_1 + \dots + m_k) = K$, and all $\theta \in \mathcal{M}_\mu$.

2. There are real numbers λ_j , for $1 \leq j \leq k$, that are the strong Lyapunov exponents $\lambda(\theta, x_0) = \lambda_j$, for all $(\theta, x_0) \in \mathcal{M}_\mu \times W_j(\theta)$, with $\|x_0\| = 1$, and one has $\lambda_k < \dots < \lambda_1$.
3. The Ergodic Spectrum $\text{LY}\Sigma(\mu)$, which depends on the ergodic measure μ , is this collection $\{\lambda_1, \dots, \lambda_k\}$. The Lyapunov Spectrum $\text{LY}\Sigma$ is the union

$$\text{LY}\Sigma \stackrel{\text{def}}{=} \bigcup_{\mu} \text{LY}\Sigma(\mu),$$

over all ergodic measures μ , is used below.

4. For each j , the measurable vector bundle

$$\mathcal{M}_\mu \times W_j(\theta) = \{(\theta, x_0) \in \mathcal{M}_\mu \times W_j(\theta)\}$$

is an invariant set for the projective flow. Furthermore, because of the exponential separation between these vector bundles, the bundle $\mathcal{M}_\mu \times W_j(\theta)$, with $j = 1$, is an attractor for the projective flow.

Connections between the $\text{SS}\Sigma$ and $\text{LY}\Sigma$: As is shown in JPS87, the following relations hold:

- One has $\text{LY}\Sigma \subset \text{SS}\Sigma$.
- For each i with $1 \leq i \leq \ell$, there is an ergodic measure μ with support in the spectral interval $[a_i, b_i]$.
- Assume that the invariant set \mathcal{M} is “dynamically connected”, that is, \mathcal{M} cannot be written as the union of two disjoint, nonempty, closed, invariant sets. Then the following holds: For b_1 , the largest value in $\text{SS}\Sigma$, there is an ergodic measure μ with the property that $\lambda = b_1$ is a strong Lyapunov exponent. It follows that $\lambda_1 = b_1$ is a Lyapunov exponent for some ergodic measure μ on \mathcal{M} , and consequently there is a unit vector x_1 in $W_1(\theta)$ with the property that x_1 is a red vector.
- The previous item is valid for any of the endpoints $\{a_1, \dots, a_k; b_1, \dots, b_k\}$, but the related ergodic measures may differ.

It can happen, as in the case with the CY92 model, that there is a unique red vector. Furthermore, in this case, as is noted in Section 4, one has $\dim V_1(\theta) = 1$. Hence one has $W_1(\theta) = V_1(\theta)$, and there is a unique red vector in the projective flow. Moreover, due to the exponential dichotomies occurring in the CY92 model, the red vector is robust, and it varies continuously with small changes in the model.

On the other hand, when $\dim W_1(\theta) = 1$, there is always a unique red vector. If in addition, one has $\dim V_1(\theta) \geq 2$, then the red vector may be only measurable and not continuous. In short, the red vector need not be robust.

3 Continuum Limits of the BV and EBV Algorithms

We now elucidate further the relationships between the BV and EBV algorithms and the dynamics of the underlying system. In particular, we now formalize the connections between these algorithms and the solutions of the linear tangent equation, i.e., the finite-time Lyapunov vectors. We are interested in the behavior as the step size δt goes to 0. To accomplish these aims we revert to a continuum formulation and for simplicity, assume that the vector $y(t)$ in (1) is a member of the Euclidean space \mathbb{R}^K . Since energy norms are used in many geophysical fluid mechanics problems, we will take the usual l^2 -norm. We denote the norm and inner product by $\|\cdot\|$ and (\cdot, \cdot) , respectively, with $\|\cdot\|^2 = (\cdot, \cdot)$

The notation in this chapter differs from what has been set in the rest of the paper in minor ways, like the letters representing the functions. This is so to emphasize that unlike what one computes in practice, the dynamical systems here are continuous in time. However, everything is clearly explained to avoid ambiguities without cluttering the presentations with technical details.

We stress that the term *continuum limit* refers to the rescaling time (sometimes called a cycle), not to the numerical integration time step. In a numerical context, this corresponds to a strategy where both integration time steps and the rescaling times are small. While this has no importance for linear systems, it leads to different outcomes when applied to nonlinear systems even for quite small perturbations amplitudes.

We first recall the system (1):

$$\begin{aligned}\frac{dy}{dt} &= G(y), \quad t > t_0, \\ y(0) &= y_0.\end{aligned}\tag{24}$$

Our first goal below is to obtain the formula

$$W(T; \varepsilon) = W(t_0; \varepsilon) + \int_{t_0}^T [G_y^\varepsilon(W(t; \varepsilon), t) - (G_y^\varepsilon(W(t; \varepsilon), t), W(t; \varepsilon)) W(t; \varepsilon)] dt,\tag{25}$$

where for all $t \geq t_0$,

$$W(t_0; \varepsilon) = W_0; \quad \|W_0\| = 1; \quad G_y^\varepsilon(W(t; \varepsilon), t) = \frac{1}{\varepsilon} G_y(\varepsilon W(t; \varepsilon), t),$$

and $G_y(\delta y, t) = G(y + \delta y) - G(y)$ for the limiting case $\delta t \rightarrow 0$, $t_0 \leq t \leq T$. The vector $W(t; \varepsilon)$ corresponds to a continuously rescaled bred vector $V(t)$ at amplitude ε and with initial perturbation $V(t_0) = \varepsilon W_0$. (W has the same direction as V , but it has amplitude 1). We will streamline the presentation by skipping some of the technical details in derivation, and we will assume at the outset that G in (24) has the necessary differentiability properties to make all the mathematical steps rigorous.

Let y and $y + \delta y$ be two solutions of (24) that satisfy the initial conditions $y(t_0) = y_0$ and $(y + \delta y)(t_0) = y_0 + \delta y_0$. Then $\delta y(t)$ is a solution of

$$\frac{d}{dt} \delta y = [G(y + \delta y) - G(y)], \quad \delta y(t_0) = \delta y_0.\tag{26}$$

Even if G is autonomous, the resulting equation (26) for δy is non-autonomous. Integrating (26) for $t \geq t_0$, we obtain

$$\begin{aligned}\delta y(t_0 + \delta t) &= \left([y(t_0) + \delta y(t_0)] + \int_{t_0}^{t_0 + \delta t} G(y + \delta y) dt \right) \\ &\quad - \left(y(t_0) + \int_{t_0}^{t_0 + \delta t} G(y) dt \right).\end{aligned}\tag{27}$$

We next define $\varepsilon = \|\delta y(t_0)\|$, and a family of vectors $v(t)$ on a sphere centered at zero with radius ε in \mathbb{R}^K by

$$v(t) = \frac{\delta y(t)}{\|\delta y(t)\|} \|\delta y(0)\|.$$

Assuming that δy is bounded away from 0, $v(t)$ obeys the evolution equation:

$$\begin{aligned} \frac{1}{\|\delta y(0)\|} \frac{dv}{dt} &= \frac{1}{\|\delta y(t)\|^2} \left[\left(\frac{d}{dt} \delta y \right) \|\delta y(t)\| - \delta y(t) \frac{d}{dt} \|\delta y(t)\| \right] \\ &= \frac{1}{\|\delta y(t)\|^2} \left[G_y(\delta y, t) \|\delta y(t)\| - \frac{1}{2} \frac{\delta y(t)}{\|\delta y(t)\|} \cdot \frac{d}{dt} \|\delta y\|^2 \right], \end{aligned} \quad (28)$$

which makes explicit the fundamental dependence of the time evolution of v on the norm. Taking the scalar product of (26) with $\delta y(t)$, we also have

$$\frac{1}{2} \frac{d}{dt} \|\delta y\|^2 = (G_y(\delta y, t), \delta y(t)).$$

Substituting this relation in (28) gives after some simplifications,

$$\frac{1}{\|\delta y(0)\|} \frac{dv}{dt} = \frac{G_y(\delta y, t)}{\|\delta y(t)\|} - \left(\frac{G_y(\delta y, t)}{\|\delta y(t)\|}, \frac{v(t)}{\|\delta y(0)\|} \right) \frac{v(t)}{\|\delta y(0)\|}. \quad (29)$$

We integrate (29) independently on successive intervals

$$[t_0, t_0 + \delta t), [t_0 + \delta t, t_0 + 2\delta t), \dots, [t_0 + (N-1)\delta t, T)$$

, where $T = t_0 + N\delta t$ and denote the solution of (29) on the interval $I_k = [t_0 + (k-1)\delta t, t_0 + k\delta t)$ by $v_k(t)$. This is a system of N integral equations with N free parameters (the integration constants) $v_k^0, k = 1, \dots, N$:

$$\begin{aligned} v_k(t) &= v_k^0 + \|v_k^0\| \int_{(k-1)\delta t}^t \left(\frac{G_y(\delta y_k, \tau)}{\|\delta y_k(\tau)\|} - \left(\frac{G_y(\delta y_k, \tau)}{\|\delta y_k(\tau)\|}, \frac{v_k(\tau)}{\|v_k^0\|} \right) \frac{v_k(\tau)}{\|v_k^0\|} \right) d\tau, \\ t_0 + (k-1)\delta t &\leq t < t_0 + k\delta t, \quad k = 1, \dots, N, \end{aligned} \quad (30)$$

where δy_k satisfies the (27) with the initial condition v_k^0 at $t = t_0 + (k-1)\delta t$, i.e., the left-hand boundary of the interval I_k . We observe that, at $t = t_0 + (k-1)\delta t$, $v_k(t)$ is an approximation of the discrete bred vector with short rescaling time step δt . Hence, we take it as the basic approximation for the continuously rescaled bred vector we seek. Note that the integrand in (30) is simply the component of the vector $\frac{G_y(\delta y, t)}{\|\delta y_k(t)\|}$ perpendicular to $v_k(t)$ in \mathbb{R}^K at time t .

We take

$$v_1^0 = \delta y(t_0), \quad v_k^0 = \lim_{t \rightarrow t_0 + (k-1)\delta t^-} v_{k-1}(t), \quad 1 \leq k \leq N, \quad (31)$$

assuming the limits exists, and extend the domain of definition of v_k by setting $v_k(t) = 0$ if $t \notin I_k$. Similarly, we extend the domain of definition of δy_k . We then let $V_N = v_1 + \cdots + v_N$, and note that V_N is continuous in time on $[t_0, t_0 + N\delta t]$ (extended to $t = t_0 + N\delta t$ by continuity) and piecewise differentiable. We define δY_N analogously as the sum of extended δy_k 's. Then

$$V = \lim_{N \rightarrow \infty} V_N = \lim_{\delta t \rightarrow 0} V_N$$

exists and is continuous in time, for instance, when G has the necessary differentiability properties so that the sequence V_N , possibly after passing to a subsequence, converges uniformly on compact time intervals, as $N \rightarrow \infty$ (or, equivalently, as $\delta t \rightarrow 0$). Since V_N coincides with the right-hand side limit of δY_N at every grid point, the difference between V_N and δY_N goes to zero as $N \rightarrow \infty$ in an appropriate norm that is allowed by how smooth G is. When G is continuously differentiable, for example, this convergence would be uniform on compact time intervals. Thus, by passing to the limit $N \rightarrow \infty$ in (30), we obtain the following representation formula for V :

$$V = V(t_0) + \int_{t_0}^T \left(G_y(V(t), t) - \left(G_y(V(t), t), \frac{V(t)}{\|V(t)\|} \right) \frac{V(t)}{\|V(t)\|} \right) dt. \quad (32)$$

Under the regularity assumptions above on G , the integrand is continuous and hence V is a mild solution of an associated differential equations. Therefore, we can bootstrap and prove the further regularity of V . Note that, by construction,

$$\varepsilon = \|V(t_0)\| = \|V(t)\|.$$

Therefore, Equation (32) is equivalent to (25), if we take $\delta y_0 = \varepsilon W_0$.

An immediate consequence of (25) is the following. When G is sufficiently regular, $W(t; \varepsilon)$ converges uniformly on $[t_0, T]$ as $\varepsilon \rightarrow 0$. Let us denote the pointwise limit by $W(t)$. Then, by (25), and the observation

$$\lim_{\varepsilon \rightarrow 0} G_y^\varepsilon(W(t), t) = \lim_{\varepsilon \rightarrow 0} \frac{G(y + \varepsilon W(t)) - G(y)}{\varepsilon} = D_{W(t)} G(y),$$

where the right-hand-side is the directional derivative of G in the $W(t)$ -direction, it follows that $W(t)$ satisfies the differential equation

$$\frac{dW}{dt} = A(t)W - (A(t)W, W)W, \quad (33a)$$

where

$$A(t) = D_y G|_{y=y(t)},$$

and

$$W(t_0) = W_0, \quad W \in \mathbb{R}^K. \quad (33b)$$

We recall the Linear Tangent Equation for the problem (24):

$$\frac{dX}{dt} = A(t)X, \quad A(t) = D_y G|_{y=y(t)}, \quad (34a)$$

$$X(t_0) = W_0, \quad X \in \mathbb{R}^K, \quad (34b)$$

where $D_y G$ is the Jacobian matrix of the map $G(y)$, and $y(t)$ is the solution of (24). As in Section 2, $U(t, s)$ denotes the solution operator of (34), taking the solution at time s to the solution and time t . With a slight abuse of notation, we write $U(t) := U(t, t_0)$, so that in particular the solution $X(t)$ of (34) at time t is simply $U(t)W_0$. We recall also that U is a semiflow, *i.e.*,

$$U(t + s) = U(t + s, t) U(t), \quad \forall t, s \in \mathbb{R}. \quad (35)$$

What is the relation between (33) and (34)? It can be easily seen that rescaling $X(t)$ to unit length and restarting the integration of (34) with $X(t)/\|X(t)\|$ at any $t_m \in [t_0, T]$ does not alter value of $X(t)/\|X(t)\|$ for $t \geq t_m$. By induction, this follows for any finite number of rescaling-restarting cycles. This is because of the linearity of (34). We leave the details to the reader. By a direct approximation argument, or by a similar argument that led to (33), we obtain

$$\frac{X(t)}{\|X(t)\|} = W(t), \quad t \in [t_0, T].$$

Thus, we have just established that, the continuously rescaled bred vector $W(t; \varepsilon)$, *i.e.*, the solution of (25), converges to the corresponding solution (rescaled to size 1) of the tangent linear equation (34) with the same initial data, *i.e.*, to the solution of (33), as $\varepsilon \rightarrow 0$ uniformly on the compact set $[t_0, T]$, assuming that G is smooth enough. Since the solutions of (33) have constant magnitude 1, it is the linear analog of (25) in terms of continuous rescaling.

While continuous rescaling is inconsequential for linear equations, this is not the case for the nonlinear ones. However, when the initial amplitude is very small, the results above indicate that one might be able to rescale less often and still get comparable results. This is simply because it takes longer for nonlinear effects to start to dominate the picture.

3.1 Continuum Limit for EBV

We discuss the continuum limit for the EBV algorithm briefly, by confining ourselves to a short, heuristic description. Writing a system of differential equation for the EBV algorithm is not as straightforward as for the BV algorithm. The perturbations $\delta y(\iota)$ in the ensemble that grow fastest at the largest amplitude (before rescaling) at time t will obey the equation (25) as long as they are the top contenders. All the other perturbations will satisfy the following differential equation:

$$\frac{d(\delta y(\iota))}{dt} = G_y(\delta y(\iota), t) - \frac{\delta y(\iota)}{\varepsilon} \frac{dh(t)}{dt}, \quad (36)$$

where $\iota \in \mathbb{I}$, and $dh(t)/dt$ is the maximum growth rate of the norm among all ensemble members with the maximum amplitude ε at time t . Assuming enough smoothness on $h(t)$, (36) can be heuristically obtained by taking the right time derivative of the vector $\delta y(\iota)(t) \frac{\varepsilon}{h(t)}$, where $h(\tau), \tau \geq t$ is the maximum norm of the ensemble members at time τ in absence of any rescaling after time t (If Equation 26 were in charge). As ε approaches zero for time t fixed, all ensemble members point in directions along which (36) approaches the linear tangent equation, since $dh(t)/dt \rightarrow 0$ by continuity of G_y at zero with $G_y(0, t) = 0$. A detailed treatment complete with the technical aspects will be presented elsewhere.

3.2 Tuning the Maximum Amplitude

We note that the convergence to the linear tangent equation is expected to be quicker for the EBV compared to the BV for most of the ensemble members. This is manifested in the equations, and later verified numerically in the next chapter. The acceleration is due to the inherent size ordering in the EBV algorithm: information on the scales where linear tangent equation dominates tends to be preserved in a robust manner against the changes in the parameter ε . On the other hand, at finite amplitude away from 0, the rescaling rule in the BV algorithm might create or prolong the life of certain instabilities due to the nonlinear effects. The algorithm output needs to be examined independently to check whether these vectors are in fact relevant to the dynamics or just artefacts of the rescaling strategy. We believe that the EBV algorithm is more resistant and robust in this regard. The parameter ε is tuned under different considerations in the EBV and BV algorithms.

3.3 Separation of Scales

In nonlinear systems, it is possible that a perturbation will not grow very rapidly in the zone of perturbations with size near ε , but instead, a multiple

of this perturbation can be dominant among the smaller perturbations and might grow quite fast there. It is also possible that the dominant vectors in different amplitude zones will be close in structure, but very different in growth characteristics. These theoretical considerations are realized in the Lorenz63 system quite strikingly. See Figure 4 and its discussion, especially the recurring patterns of “spears” therein. This can be seen very easily from the following relation:

$$\frac{\|V(t_n)\|}{\|V(t_{n-1})\|} = \mathbf{R}_n^{\min} \frac{\|\bar{V}(t_n)\|}{\|V(t_{n-1})\|},$$

where \mathbf{R}_n^{\min} is defined similarly to (9), $V(t)$ denotes a EBV(t) member, and $\bar{V}(t)$ denotes the same vector right before the rescaling. In a nonlinear system, this ratio can be larger than 1, for some range of perturbation sizes smaller than ε . This is the main mechanism that creates the vector zones of different magnitude with zonal growth characteristics, or a separation of scales.

Another case in point will be presented in Section 4 for the CY92 model. Even at late times and small amplitudes, we can see perturbations surviving the rescaling strategy and they resemble what one would get from the usual BV algorithm at those sizes (and the finite-time Lyapunov vectors) very closely, whereas the dominating perturbation of size ε resembles a particular BV of size ε . This example actually illustrates more. There are members of the ensemble in small magnitudes that are slow in aligning with the dominant directions.

4 Applications of the Bred Vector Algorithms

We will be comparing the BV and the EBV on two problems: The Lorenz equations, or Lorenz63 (see Lorenz (1963)), and a dissipative and forced nonlinear partial differential equation that arises in modeling the thermohaline circulation (see Cessi and Young (1992)). The latter will be denoted as the CY92. It is a Cahn-Hilliard equation and it will shown to have an inertial manifold. We will also have occasion to compare the BV and EBV results to the finite-time Lyapunov vector outcomes.

Throughout we will use an explicit fourth order Runge-Kutta time marching scheme, for the calculation of the base solutions as well as for the calculations of the BV, EBV, and finite-time Lyapunov vectors.

4.1 The Lorenz63 Model

The finite-dimensional, nonlinear Lorenz63 model has often been used as benchmark for testing sensitivity and, in particular, as a test problem for

BV (see for example Evans et al. (2004)).

Let $X = (X_1, X_2, X_3) \in \mathbb{R}^3$ and let

$$A = \begin{pmatrix} -\sigma & \sigma & 0 \\ r & -1 & 0 \\ 0 & 0 & -b \end{pmatrix}, N(X) = \begin{pmatrix} 0 \\ -X_1 X_3 \\ X_1 X_2 \end{pmatrix}, DN(X) = \frac{\partial}{\partial X} N(X). \quad (37)$$

$DN(X)$ is a 3×3 matrix-valued function. The Lorenz model is described by the solutions of the nonlinear system:

$$\partial_t X = A X + N(X). \quad (38)$$

The associated Tangent Linear Model is the skew product system

$$\begin{aligned} \partial_t X &= A X + N(X) \\ \partial_t U &= (A + DN(X)) U, \end{aligned} \quad (39)$$

where $U = (U_1, U_2, U_3) \in \mathbb{R}^3$. The U -equation in (39), which is a linear equation, is of special interest to us.

We will use the notation for the Tangent Linear Model introduced in Section 2.3. With the initial condition $\theta = X_0 \in \mathbb{R}^3$, we let $\theta \cdot t = S(\theta, t) = X(t)$ denote the solution of the nonlinear equations (38) that satisfies $S(\theta, 0) = \theta$. In this study we set $r = 28, b = 8/3, \sigma = 10$. It is well-known that for these values of the parameters the Lorenz63 model has a chaotic global attractor. For the non-autonomous U -equation; $\partial_t U = (A + DN(\theta \cdot t)) U$, we let

$$U(t) = U(\theta, t) U_0, \quad \text{denote the solution operator,} \quad (40)$$

where $U_0 \in \mathbb{R}^3$, and $U(0) = U_0$.

We will use the term “attractor” to refer to a compact, invariant set \mathfrak{A} that attracts a neighborhood of itself. As a result, \mathfrak{A} is Lyapunov stable, as a set. Consequently, for $0 < \varepsilon < 1$, there is a family of ε -neighborhoods, N_ε , of \mathfrak{A} , where each neighborhood is positively invariant and $\mathfrak{A} = \bigcap_{\varepsilon > 0} N_\varepsilon$. When we write that θ is near \mathfrak{A} , we mean that $\theta \in N_\varepsilon$, for some small $\varepsilon > 0$. See Chapter 2 in Sell and You (2002), for a history of this concept and more information.

By using a somewhat different –but equivalent formulation– Toth & Kalnay have suggested that the time evolution $BV(t)$ is a good approximation of the tangent linear solution $U(t)$, over bounded time-intervals. As a consequence of the continuum limit theory in Section 3, we see that this perceptive observation has a solid mathematical basis, provided that the time-step δt is small and the perturbation amplitude is small, as well.

In order to find a numerical validation of the continuum limit theory described in Section 3, we will calculate the distance

$$D(t, \delta t) := d_{\text{pr}}([BV(t)], [U(t)]), \quad (41)$$

where d_{pr} is the projective metric on $\mathfrak{A} \times \mathcal{P}^{K-1}$ (see (15)), and $BV(t)$ is a BV at time t . For this calculation, we assume that both $BV(t)$ and $U(t)$ satisfy the same initial condition $(y_0, \delta\mathcal{Y}_0)$ at $t = 0$, and that $y_0 = \theta$ is near the attractor \mathfrak{A} . We also use identical ensembles of perturbation vectors for both the BV and the EBV algorithms. We then fix $T > 0$, and we examine the distances $D(T, \delta t)$ (in the projective space), for different choices of δt . Our goal is to show that $D(T, \delta t)$ becomes smaller, as δt gets smaller. The max and min values, for both the BV and the EBV algorithms, are reported in Table 1. For the calculations used to generate the data in Table 1, we had set the initial conditions for both $BV(t)$ and $U(t)$, at $t = 0$, so that $X = 0.5688$, $Y = 0.4694$, $Z = 0.0119$, and $\delta\mathcal{Y}_0 = [1, 1, 0]$. For the ensemble, we took the 3-fold Cartesian product of the set

$$\{-1, -0.75, -0.5, -0.25, 0, 0.25, 0.5, 0.75, 1\}$$

in \mathbb{R}^1 , and projected this product onto the sphere of radius 0.1, centered at the origin, in \mathbb{R}^3 . By eliminating the repetitions introduced with this projection, one obtains an ensemble of 584 distinct points in \mathbb{R}^3 . (Note that the initial conditions are "near" the Lorenz attractor.) The trajectory for the nonlinear problem was computed with a step size $dt = 0.0001$, so that any observed variations are only due to the differences in the algorithms and the relative effects on the perturbations. In Table 1, we fixed $T = 2$ and made two choices for δt , namely $\delta t = 0.004$ and 0.001 . The results compare the outcomes of the BV and the EBV algorithms. As noted in Section 3, both algorithms $EBV(t)$ and $BV(t)$ approximate solutions of the tangent linear equation, as δt goes to 0. By using the projective metric one can compare the rates of convergence in terms of this metric.

The maximal values of $D(T, \delta t)$, for both the BV and EBV algorithms, are essentially the same for both choices of δt , and the minimal values are essentially the same for $\delta t = 0.004$. When one moves from $\delta t = 0.004$ to $\delta t = 0.001$, both minimal values decrease, as expected. However the drop in the minimal value for the EBV algorithm is substantially larger than the drop for the BV algorithm. By using the perturbation corresponding to the minimal drop in $D(T, \delta t)$, for each of the algorithms, one arrives at the *best approximation* given by for the given algorithm. Clearly the min BV and min EBV columns in Table 1 shows that the EBV algorithm yields a better approximation than the BV algorithm.

One can, of course, use different initial conditions for the two solutions $BV(t)$ and $U(t)$ and/or larger perturbations. However, one cannot expect to replicate the results seen in Table 1 in that case. First, there is a transient phase, which ends when the two solutions are "near" the attractor. Even if this transient phase is short, one still has a problem. While the attractor is stable, as a set, one still has a problem because the flow on the attractor

Table 1: *Maximum and minimum values of $D(T, \delta t)$ with $T = 2$, for two δt . See (41).*

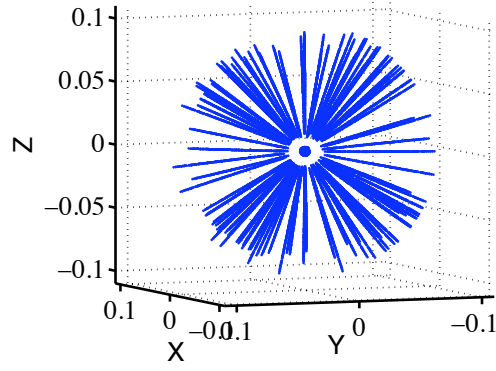
δt	max BV	min BV	max EBV	min EBV
0.004	1.40×10^{-1}	5.26×10^{-4}	1.01×10^{-1}	1.10×10^{-3}
0.001	7.77×10^{-2}	1.03×10^{-4}	3.48×10^{-2}	4.16×10^{-5}

is generally not Lyapunov stable. Essentially all pairs of nearby orbits may diverge over long time intervals.

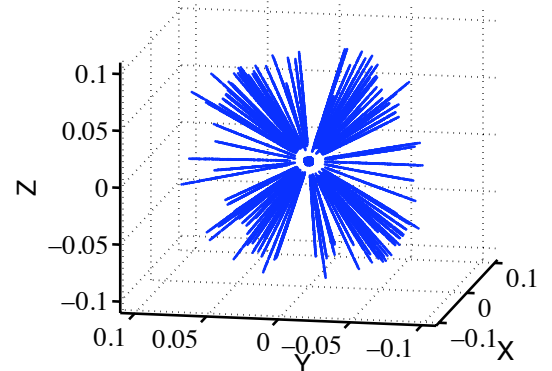
An interesting outcome of the use of the EBV algorithm on Lorenz63 is that it exposes the fractal behavior of the Lorenz attractor. We computed a 584-member EBV, using the same parameter values, initial conditions and base trajectory as was used in generating Table 1, for the $\delta t = 0.001$ case. We constructed plots by merging all the vectors between time 24 and 30, at intervals of $dt = 0.1$. In that time, at any given instant, one only observes a couple of members reaching highest amplitude. By rotating the same plot about the $X_3 = Z$ axis, one obtains three views of the EBV's shown in Figure 1. Zooming into Figure 1a by a factor of 8, 32, and 60, respectively, we obtain Figure 2. The fact that one observes similar patterns at several levels of magnification meets the requirement of ‘fractal behavior’ that was introduced by Mandelbrot (1977).

Figure 3 shows the time evolution of a sample BV and the corresponding finite-time Lyapunov vector with the same initial perturbation. The results were obtained with a time step of $\delta t = 0.005$. The computations of the BVs and finite time Lyapunov vectors were performed with a time interval equal to the dynamics time step. At $t = 0$, $X = 0.1493$, $Y = 6.2575$, $Z = 1.8407$. The initial perturbation vector $\delta \mathcal{Y}_0$ was $[1, 1, 1]$. As expected, small to moderate-sized perturbations produces similar results in the finite-time Lyapunov calculation and the Bred Vector calculation.

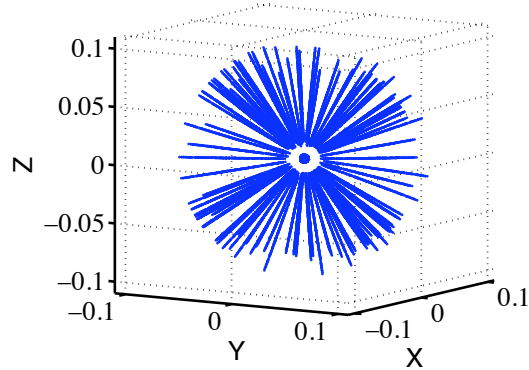
Figure 4 is devoted to the results of a calculation of $EBV(t)$ alone, for the same case. The figure depicts a rescaled time evolution of the norm of 98 distinct ensemble members in the EBV calculation. The rescaling is done with respect to the usual Euclidean l^2 -norm. The initial perturbations are made to sample a perturbation sphere of amplitude 1 about the initial conditions. The figure highlights the rapid decay of many of the vectors and the eventual size-ordering that is inherent in the EBV algorithm. For $0 < t \leq 13$, the results in Figure 4 describe the transient behavior and are rather chaotic. However, for $t \geq 13$ a very interesting pattern evolves: we see that the largest member of the ensemble takes on the value 1, for all $t \geq 0$, which is expected. (This corresponds to the Lyapunov vector with the largest exponent.) From the totality of the information on the



(a)



(b)



(c)

Figure 1: Viewed from three different angles, (a)-(c). EBVs, for times 24 through 30, taken at 0.1 time intervals, for Lorenz63. Parameters and conditions are those used to generate Table 1.

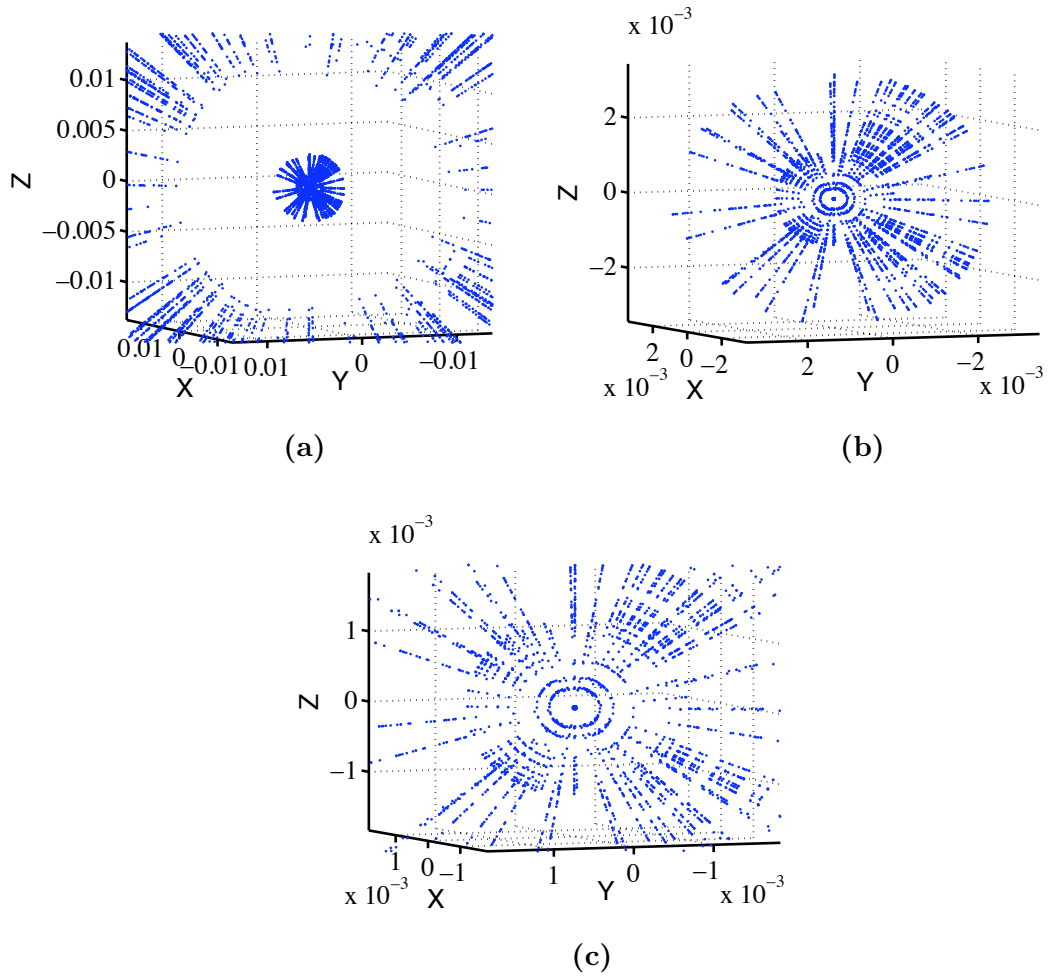


Figure 2: *Close-up views of Figure 1a: (a) zoomed in 8 times; (b) zoomed in 32 times; (c) zoomed in 60 times. Note the preservation of structure even as we zoom in.*

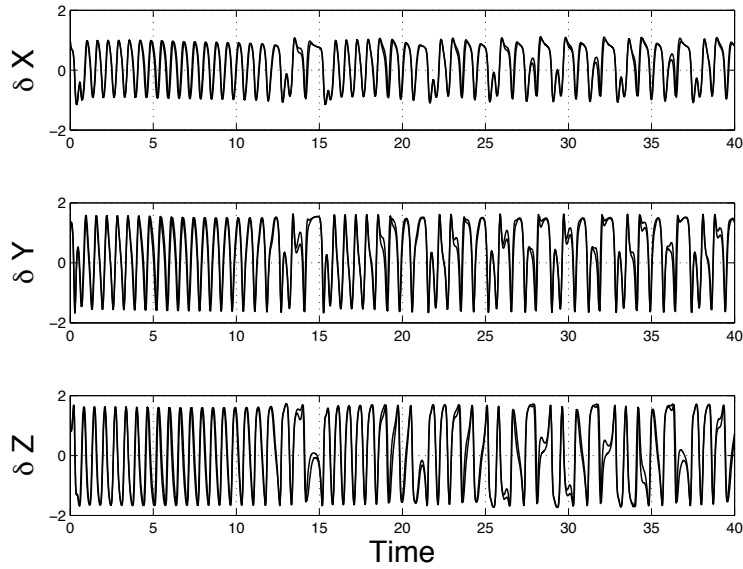


Figure 3: *The three components of BV and of the finite-time Lyapunov vectors, as a function of time, for the Lorenz63 system. (For parameters and initial conditions, see text). The vectors are nearly coincidental over the whole time span. The perturbation size is 1, in all of the components. For our choice of initial conditions the vectors exhibit more temporal regularity at earlier times. This transient behavior disappears at $t = 13$, approximately.*

ensemble, it is possible to extract the next two Lyapunov vectors, where the exponents satisfy $\lambda_3 \leq \lambda_2 \leq \lambda_1$. However, this would require a deeper analysis, and it is deferred to a future work. In fact, viewed as a whole, the graphs corresponding to those smaller than 1 have very useful structural information: Notice the recurrent ‘spear-like’ pattern, which occurs after $t = 17$. What is happening is that a group of ‘small’ vectors in the attractor grow rapidly, as they go around the horn in the Lorenz attractor. We propose that the rationale for this behavior is that the nonlinear equations of motion temporarily overwhelm the uniform rescaling rule for these small vectors. For $t \geq 17$, we are beyond the transient zone, and we would not expect such behavior to occur if the linear term AX strongly dominates the nonlinear term $N(X)$ in the vicinity of the attractor. This is an excellent illustration of the fact that EBV algorithm preserves the role of the nonlinear terms in the equations of motion.

This feature of the EBV can not be replicated by BV, even a BV with an ensemble of perturbations. The reason for this is that the rescaling rule for the BV algorithm forces all the perturbations to have the same norm, for each $t \geq 0$. Thus the Figure 4 for the EBV would be replaced by a figure for the BV, where all of the perturbations are plotted on the top line only.

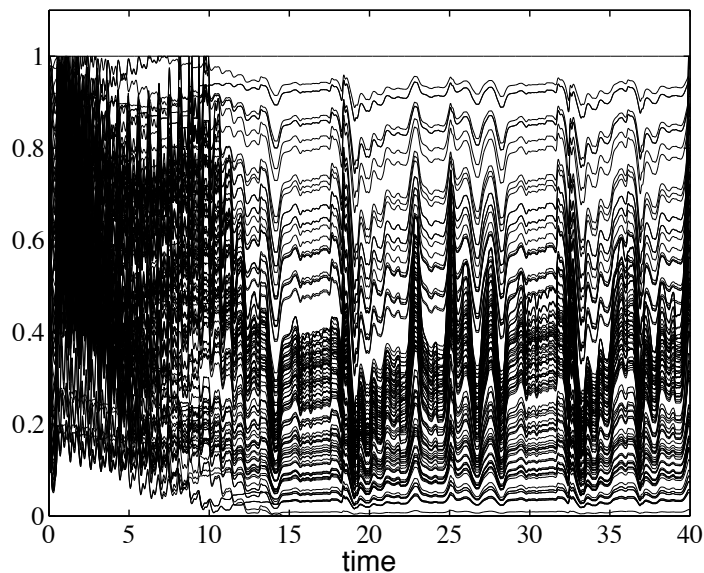


Figure 4: *Plot depicting the evolution of the 2-norm of an ensemble of 98 $EBV(t)$. Same parameters and initial conditions, as in Figure 3. Initially, in the interval $[0, 13]$ we note a very fast decay of some of the ensemble members, leading to a sorting in size, beyond that time. As described in the text, this outcome is one of the most distinguishing features of the EBVs, when compared to an ensemble of individual BV outcomes.*

4.2 A Cahn-Hilliard Equation

In their work on the thermohaline dynamics Cessi and Young (1992) proposed a coupled model for the circulation, salinity, temperature, and density of the oceans, with atmospheric forcing. The crux of the model is the partial differential equation for salinity: The slow-time dynamics of the ocean salinity $S(x, t)$, zonally-averaged, and as a function of latitude $x \in [-\pi, \pi]$ and time $t \geq 0$, is described by

$$\begin{aligned} \frac{\partial S}{\partial t} &= \alpha \frac{\partial^2}{\partial x^2} [f(x) + \mu S(S - \sin(x))^2 + S - \gamma \frac{\partial^2 S}{\partial x^2}], \quad t > 0, \\ S(x, 0) &= S_0(x). \end{aligned} \quad (42)$$

The equation is subject to zero-flux and zero-stress boundary conditions at the poles, however, we will be considering periodic boundary conditions (for steady and periodic forcing $f(x)$ as well as equilibrium solution $\sin(x)$ the conclusions that follow apply to the zero-flux case). The positive parameter α affects the strength of the linear stability of the model. We fix $\alpha = 3.5 \times 10^{-3}$, $\gamma = 0.001$, and $\mu = \sqrt{10}$. The forcing $f(x)$ is a prescribed function that reflects balances of evaporation and precipitation of freshwater; it can be symmetric, about the Equator ($x = 0$), but is more typically, asymmetric (see Eyink (2005)). The second derivative of the forcing function, will be chosen to be

$$\partial_{xx} f(x) = \begin{cases} \frac{2}{27}(36 - 39\mu^2 \cos^2 x - 81 \cos^2 x + 8\mu^2 + 25\mu^2 \cos^4 x) \sin x, \\ \quad \text{for } x \in [-\pi, 0], \\ \\ -\frac{2}{9}(3\mu^2 \cos^2 x - 3 - \mu^2) \sin x, \\ \quad \text{for } x \in (0, \pi]. \end{cases} \quad (43)$$

It is shown in Figure 5a. The initial condition chosen for this computation is $S_0(x) = \cos(x)$. The base solution obtained numerically is displayed in Figure 5b.

A great deal is known and can be said about the mathematical structure of CY92 and its solutions. By letting $S = \partial_x u$ and $f = \partial_x F$ the CY92 can be related to the Cahn-Hilliard Equation (CHE) with forcing: The general CHE, for u above, is

$$\partial_t u + \nu \Delta^2 u = \Delta(g(u)) + F \quad \text{for } x \in \Omega \text{ and } t \geq 0, \quad (44)$$

where $u = u(t, x)$ is a scalar field, $\nu > 0$ is a constant, Δ is the Laplacian operator, Δ^2 is the bi-harmonic operator, and g is a polynomial of degree 3: $g(u) = \sum_{j=1}^3 a_j u^j$, with $a_3 > 0$. The term $F = F(x)$ is the forcing function. In general, the domain Ω may be an open bounded domain in the Euclidean space \mathbb{R}^m , with $1 \leq m \leq 3$. However, we restrict our attention, to the case

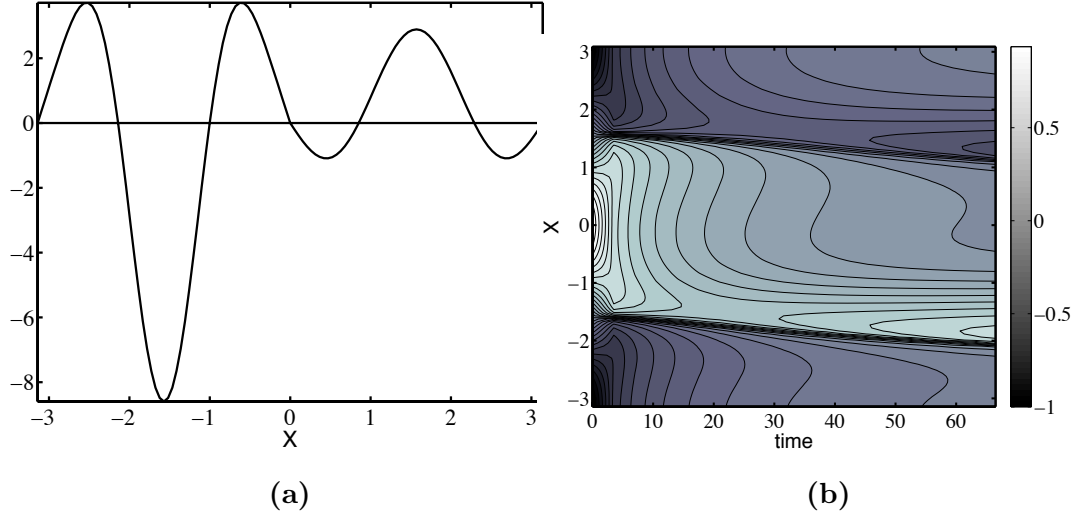


Figure 5: (a) Forcing function f_{xx} , (43), used in the CY92 model; the numerical solution to CY92 is shown in (b). See Figure 6 for the BV algorithm and finite-time Lyapunov vector algorithm outcomes.

where $m = 1$ and Ω is the interval $\Omega = (-\pi, \pi)$, with boundary $\Gamma = \{\pm\pi\}$. As we will see, the multiplicity of the largest Lyapunov exponent for this problem is 1.

For the analysis of the solutions of the CHE, one will use the standard Sobolev spaces $H^j = H^j(\Omega)$, where $j \geq 0$ is an integer and $H^0 = L^2(\Omega)$. As usual, the inner product and norm on H^0 is denoted by $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_0$ and $\| \cdot \| = \| \cdot \|_0$. For $j \geq 0$ one uses

$$\|u\|_{j+1}^2 = \|u\|_j^2 + \sum_{|\alpha|=j+1} \int_{\Omega} |D^\alpha u|^2 dx,$$

where

$$D^\alpha = \frac{\partial^\alpha}{\partial x^\alpha}, \quad \alpha \geq 0,$$

and

$$\langle u, v \rangle_{j+1} = \langle u, v \rangle_j + \sum_{|\alpha|=j+1} \int_{\Omega} (D^\alpha u) (D^\alpha v) dx.$$

The finite-time Lyapunov vector algorithm for the CY92 model is derived, by first rewriting the equation as

$$\partial_t S = \alpha \partial_{xx} [\mu^2 S (S - \eta)^2 - r f(x) + S - \gamma^2 \partial_{xx} S] =: F(S).$$

We opt here to first linearize and then discretize. To obtain the tangent linear equation, we need to find the linear map \mathcal{L} such that

$$F(S + \delta S) - F(S) = \mathcal{L}(\delta S) + o(\delta S) \text{ as } \|\delta S\| \rightarrow 0,$$

where $\|\cdot\|$ is a norm. Typically, this will be the norm in the space where solutions live, such as the Sobolev space H^1 for S , or dictated by physical considerations. In this infinite-dimensional model, different norms are not necessarily equivalent.

First, we note that

$$\{S + \delta S\} (\{S + \delta S\} - \eta)^2 - S (S - \eta)^2 = (S - \eta) (3S - \eta) + o(\|\delta S\|).$$

We let L be a finite-difference approximation to \mathcal{L} . The discrete tangent linear equation becomes the product of two matrices AB , where B is the matrix whose entries are found by discretizing the operator

$$\mu^2 (S - \eta) (3S - \eta) + Id - \gamma^2 \partial_{xx},$$

where Id is the identity matrix, and A is the matrix corresponding to the discretization of $\alpha \partial_{xx}$. The finite-time Lyapunov algorithm is obtained by solving

$$\partial_t \delta \Theta = L \delta \Theta = AB \delta \Theta = A(B(\delta \Theta)),$$

for $t \in [0, T]$, subject to some initial vector perturbation $\delta \Theta(0) = s$, where $\delta \Theta$ is the discretization of the perturbation δS . The perturbation is normalized to the norm of the initial perturbation at each δt , the time step of the explicit Runge-Kutta 4 time integration scheme employed here. $T = 70$. In the simulations, $\delta t = 0.01$. We applied second-order centered finite differences in space, and used 121 grid points, $-\pi = x_0, x_1, \dots, x_{120} = \pi$.

By construction the finite-time Lyapunov vectors are not amplitude sensitive, however the BVs are, thus different amplitude perturbations will yield different BVs, in the case of a general nonlinear problem. This outcome has important practical implications, if one would like to use BV to either infer the structure in the field, or the degree of sensitivity of the outcomes to perturbations in initial conditions. A challenging problem could thus arise in the context of large-scale simulations: what is considered a large structural change or a highly sensitive outcome is physics-dependent, perhaps even difficult to surmise quantitatively; the physics in question may not be fully understood and thus a reasonable perturbation amplitude is simply guessed.

We ran the finite-time Lyapunov vector and the BV algorithms, using the same initial condition, forcing and perturbation. First, we examine the effect of the size of the amplitude of the perturbation on the outcomes: we do so by keeping the shape of the perturbation fixed, changing only the overall

amplitude. Figure 6a and b are plots of the final BVs and finite-time Lyapunov vectors, corresponding to 2-norm 0.25 and 0.025 sized perturbations, respectively. When the perturbations are small the BVs and the finite-time Lyapunov vectors are qualitatively consistent and, nearly so, quantitatively. The outcomes shown here are typical of the general case, that is, the qualitative and quantitative disagreement grows with an increase in the amplitude of perturbations.

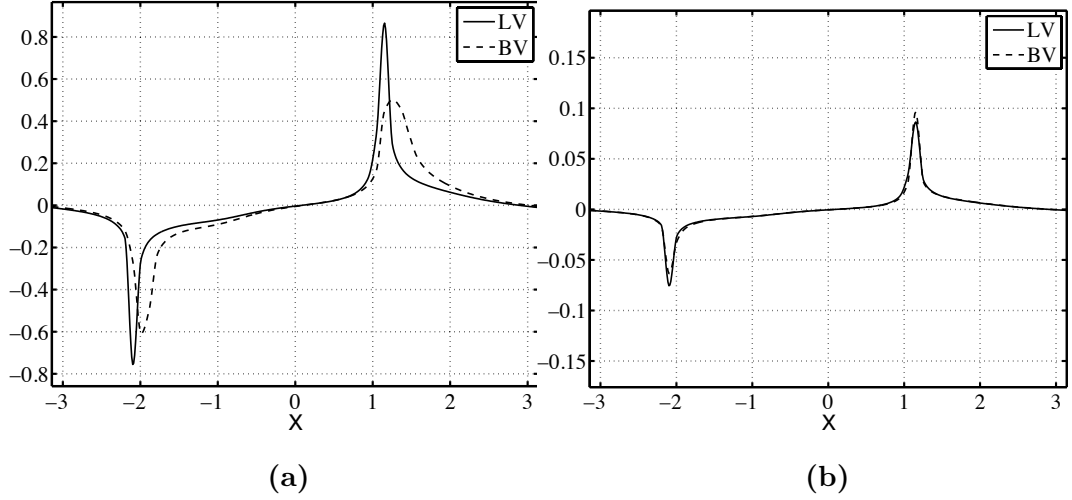


Figure 6: *Comparison of the finite time Lyapunov vectors and BVs. Effect of amplitude of perturbation on the CY92 Model, with initial conditions $Y_0 = \cos(x)$. The initial perturbation was $\epsilon \sin(x)$. Comparison of the finite-time Lyapunov vector and the BV outcomes at $t = 70$. (a) Corresponds to $\epsilon = 0.25$; (b) to $\epsilon = 0.025$.*

The shape or spectral content of the perturbation mattered as well. The spectrum of the perturbation is clearly important when projecting onto spectral bases for a reduced representation. To illustrate this, we use the CY92 model, with the same forcing as before and same initial condition. We examine monochromatic sine wave perturbations with wavenumber $j = 1, 2, \dots, 6$ of the form

$$\delta Y_0 = \epsilon_j \sin[jx + \frac{1}{3}(j-1)\exp(1)], \quad (45)$$

where all of the perturbation amplitudes remain the same, $\epsilon_j = 0.25$. The choice of the phase is inconsequential: to show this we chose non-commensurate

phases among the sine wave components. In Figure 7a we show the BVs associated with each of the perturbations ($j = 1, 2, \dots, 6$), at $t = 70$. Figure 7b shows the EBVs at $t = 70$. Even at $t = 5$, shown in Figure 7c, we already see an amplitude-ordered structure in the EBV.

The space-time plot of all EBVs is plotted in Figure 7d; we note the very short transient phase, lasting till about $t = 4$, followed by a structure which is clearly dominated by the largest member of the EBV. The EBV calculation was run with identical parameters to those used in Figure 7a, with the ensemble consisting of the same initial perturbations in (45). In the BV case we are getting outcomes that do not have the reductive appearance of the EBV. On the other hand, there is no ambiguity to the prevailing ensemble member in the EBV case. This ensemble member has a clear correspondence in structure to the path and its perturbation field. The BVs should eventually agree with the EBVs, nevertheless. It is safe to assume that this will happen only after a very long time, longer than the time interval that might be suggested by the base solution.

We also compared BV, for each of the perturbations in (45), to the outcomes of the finite-time Lyapunov vector calculation. Figure 8 compares the BVs and finite-time Lyapunov vectors for each j wavenumber perturbation. The amplitudes are set to $\epsilon_j = 0.25$, for $j = 1, 2, \dots, 6$. The structure of the finite-time Lyapunov vectors, to within a sign, is qualitatively the same, regardless of the wavenumber of the perturbation. The BVs are not. If the perturbation was made considerably smaller the differences between the BVs and the finite-time Lyapunov vectors would become small, as expected: As shown in Section 3, the BVs and the finite-time Lyapunov vectors must be similar to each other, provided the perturbations are small enough. For larger amplitudes, the BVs might show more structure since nonlinear effects could play a role in the structure of the BVs. Apparently, $\epsilon = 0.25$ is already in the range of large perturbations of the CY92 about the solution chosen. It was not clear whether a chosen perturbation is large or small, based solely on the CY92 model itself: it was only clear to us after a comparison of the outcomes of the finite-time Lyapunov case and the BV case.

4.3 Cahn-Hilliard Dynamics

The Bi-harmonic Operator $Bu = \Delta^2 u$ on Ω and the linear Bi-harmonic equation:

$$\partial_t u + \nu \Delta^2 u = 0 \quad (46)$$

play a basic role in the study of the CHE. We assume for now that the Operator B satisfies either the non-flux boundary conditions:

$$\frac{\partial u}{\partial n} = 0 \quad \text{and} \quad \frac{\partial(\Delta u)}{\partial n} = 0, \quad \text{on } \Gamma, \quad (47)$$

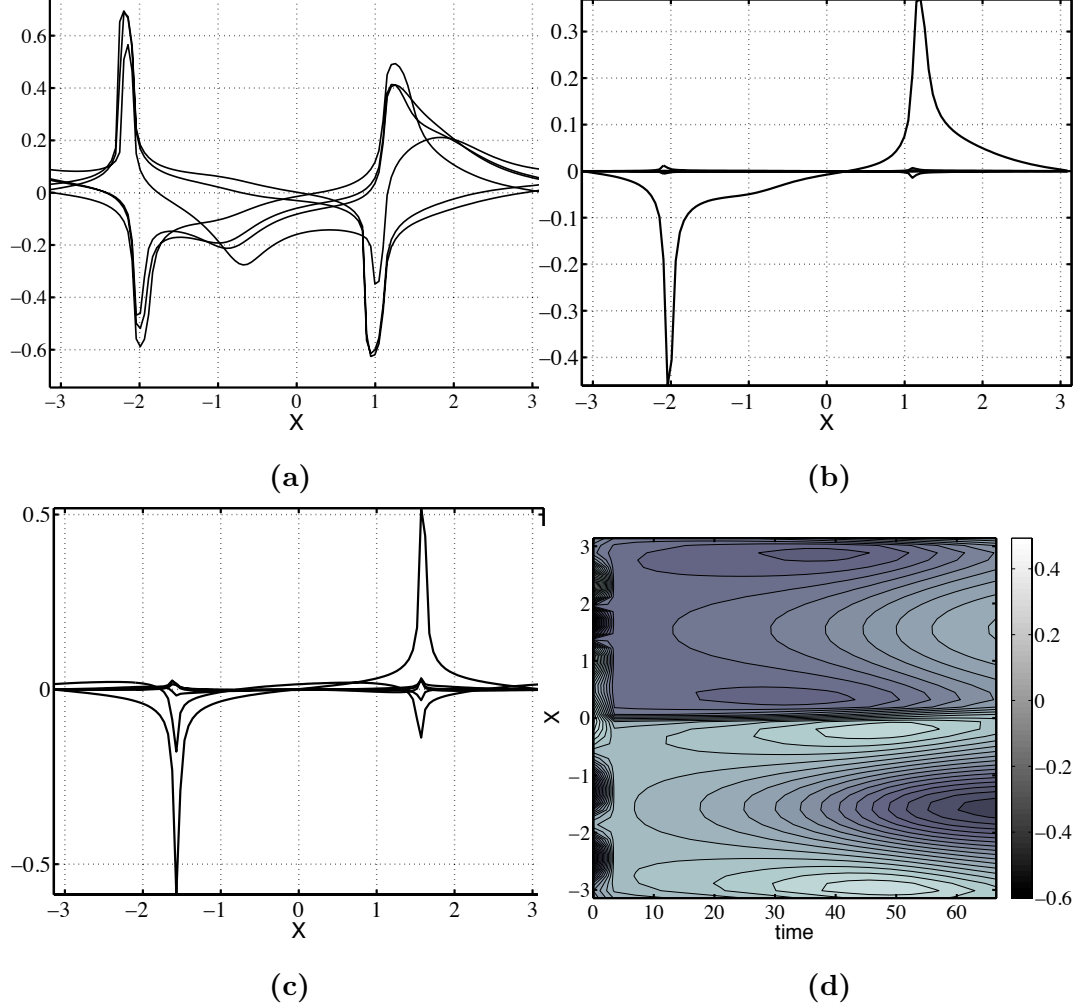


Figure 7: (a) At $t = 70$, the superposition of 6 BV simulations, the j^{th} one corresponding to a sine wave perturbation, as per (45). (b) Cross section of the EBVs, at $t = 70$, run with an ensemble of perturbations (see 45). We note that, unlike the BV case in (a), the EBV has settled into a structure that clearly reflects the dominant EBV ensemble member. (c) Cross section of the same EBVs, at an earlier time: at $t = 5$. Even for short times, the EBV already shows an amplitude-ordering of its vectors. (d) Superposition of all EBVs, as a function of time and space.

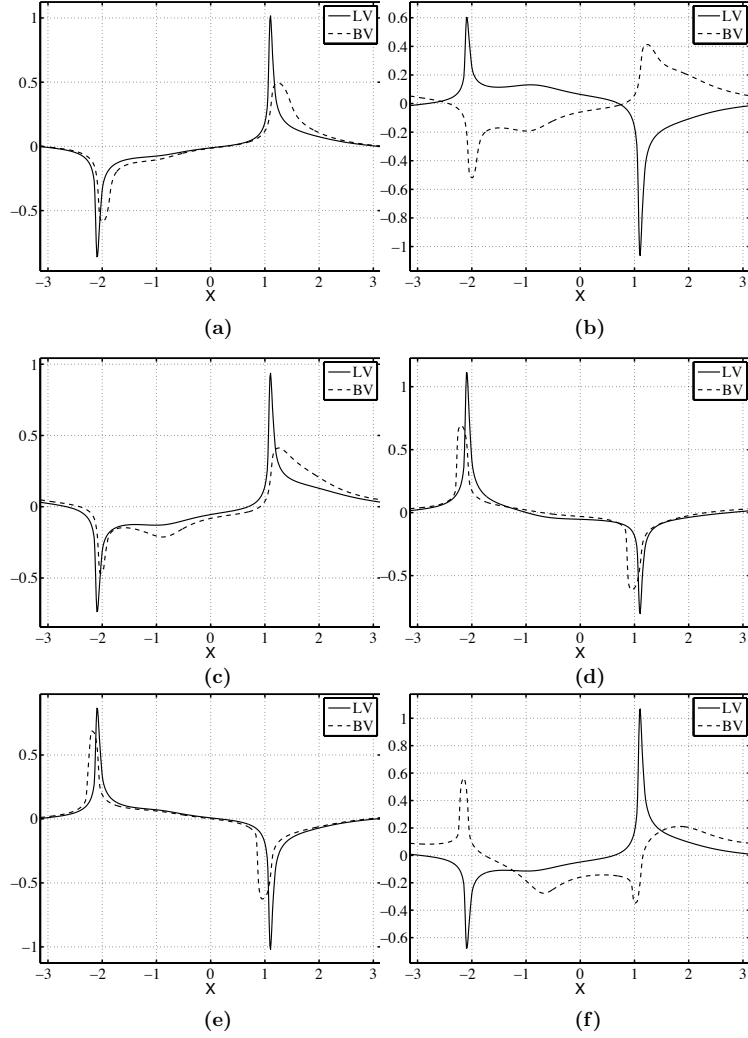


Figure 8: Wavenumber dependence of BVs and finite-time Lyapunov vectors. CY92 Model, cosine initial condition and non-symmetric forcing (see Figure 5). Comparison of the finite-time Lyapunov vectors and the BVs for sine wave perturbations of size $\epsilon_j = 0.25$. The perturbations are the individual sine waves in (45), (a) $j = 1$, (b) $j = 2$, ..., (f) $j = 6$.

or the periodic boundary conditions, see Sell and You (2002). In the sequel, we will assume that the forcing function F in (44) satisfies $F \in H_0^j$, for some $j \geq 0$. That is to say,

$$\int_{\Omega} F \, dx = 0.$$

Thus, by integrating (44), one observes that any solution $u = u(t, x)$ satisfies

$$\bar{u}_0 := \frac{1}{|\Omega|} \int_{\Omega} u_0(x) \, dx = \frac{1}{|\Omega|} \int_{\Omega} u(t, x) \, dx = \bar{u}(t), \quad \text{for } t \geq 0. \quad (48)$$

One seeks solutions of the CHE (44) in the Sobolev space V_0^2 . A *mild* solution $u = u(t)$ of the initial value problem is given by

$$u(t) = e^{-\nu B_0 t} u_0 + \int_0^t e^{-\nu B_0(t-s)} [\Delta(g(u(s))) + F] \, ds. \quad (49)$$

We will denote the maximally defined solutions of (49) by $u(t) = S(t)u_0$. With $u_0 \in V_0^2$, this mild solution is uniquely determined, with $S(t)u_0 \in V_0^2$, for $0 \leq t < T(u_0)$, where $0 < T(u_0) \leq \infty$.

Because of (48), we see that, for $j \geq 2$, the spaces

$$H_0^j := \{\phi \in H^j : \bar{\phi} = 0\} \quad (50)$$

are positively invariant spaces for the solutions of (44). Thus we will focus only on solutions $S(t)u_0$ that satisfy $\bar{u}_0 = 0$. We also denote the collection of stationary solutions of the CHE by

$$Q := \{u_0 \in H^2 : S(t)u_0 = u_0 \text{ for all } t \in \mathbb{R}\},$$

and $Q_0 = Q \cap H_0^2 = Q \cap V_0^2$.

4.3.1 The CHE With $F \equiv 0$:

The basic problem with forcing $F \equiv 0$ is of special interest. The equation (44) becomes:

$$\partial_t u + \nu \Delta^2 u = \Delta(g(u)) \quad \text{for } x \in \Omega \text{ and } t \geq 0. \quad (51)$$

To study the solutions of (51), one uses the Landau-Ginsburg functional:

$$J(u) := \int_{\Omega} \left[\frac{\nu}{2} |\nabla u|^2 + G(u) \right] dx, \quad \text{where } G(z) = \int_0^z g(s) \, ds. \quad (52)$$

In addition, the Landau-Ginsburg functional satisfies

$$\partial_t J(S(t)u_0) = -\|\nabla K(S(t)u_0)\|_0^2, \quad \text{for } 0 < t < T(u_0), \quad (53)$$

where $K(u) = -\nu\Delta u + g(u)$. Furthermore, there exist positive constants f_0 and C_0 such that $-f_0 \leq G(s) \leq C_0 s^4 + f_0$ and

$$\frac{\nu}{2} \|\nabla S(t)u_0\|_0^2 - f_0|\Omega| \leq J(u_0). \quad (54)$$

Since $J(u_0)$ is bounded below, see (54), and decreasing along orbits, see (53), it follows that $J(S(t)u_0)$ is a Lyapunov function, see LaSalle and Lefschetz (1961). This implies that the mild solution $S(t)u_0$ is defined for all $t > 0$, i.e., $T(u_0) = \infty$, for every $u_0 \in V^2$. It is fact that, whenever $F \equiv 0$ and $\bar{u}_0 = 0$, then $\omega(u_0)$, the omega limit set of the solution $S(t)u_0$ is a nonempty, compact, connected invariant set in Q_0 .

Let us now return to the CHE with forcing (44). By integrating the equation (44), where $u_0 \in H_0^2$ and $F \in H_0^2$, as well, then $S(t)u_0 \in H_0^2$, for all $t \geq 0$.

4.3.2 The Global Attractor \mathfrak{A}_0 .

Assume that B satisfies the boundary conditions BC, *i.e.*, either the non-flux condition (47) holds, or the periodic boundary conditions hold. Let $S(t)$ be the semiflow generated by the solution operator on V_0^2 . Then the following hold.

1. $S(t)$ has a nonempty, compact global attractor \mathfrak{A}_0 in V_0^2 , and \mathfrak{A}_0 attracts all bounded sets in V_0^2 . The attractor \mathfrak{A}_0 depends continuously on the forcing function $F \in H_0^2$.
2. When $F \in H_0^3$, then the attractor \mathfrak{A}_0 is a compact, invariant set in V^{4r} , for each r with $0 \leq r \leq 1$.
3. The set Q_0 is nonempty, compact, and invariant with $Q_0 \subset \mathfrak{A}_0$.
4. Lastly, there is an Inertial Manifold for the solutions of the infinite dimensional system (44), see Foias et al. (1988) and Sell and You (2002). One finds this manifold by using the orthogonal projection \mathbb{P}_N onto the lowest N nodes, that is to say, into $\text{Span}\{e_k : k \leq N\}$, where $\{e_k\}$ are the eigenfunctions for the Bi-harmonic operator B . One then makes a change of variables $u = v + w$, where $v = \mathbb{P}_N u$, $w = \mathbb{Q}_N u$, and $\mathbb{Q}_N = I - \mathbb{P}_N$. The (v, w) system:

$$\begin{aligned} \partial_t v + Bv &= \mathbb{P}_N (\Delta(g(v+w)) + F) \\ \partial_t w + Bw &= \mathbb{Q}_N (\Delta(g(v+w)) + F), \end{aligned} \quad (55)$$

is equivalent to the CHE (44). One then shows that, for N large, the variable w is enslaved to the variable v in some neighborhood of \mathfrak{A}_0 .

That is to say, $w = \Phi(v)$, for a suitable function Φ . It turns out the the longtime dynamics of (44) is equivalent to the longtime dynamics of the finite-dimensional ordinary differential equation:

$$\partial_t v + Bv = \mathbb{P}_N(\Delta(g(v + \Phi(v))) + F). \quad (56)$$

See Sell and You (2002) and the references contained therein, for more details.

4.4 BV and EBV in the Higher Nonlinear Regime

For sufficiently small perturbations and short renormalizing time intervals the finite-time Lyapunov vector, the BV algorithm, and the EBV are similar in outcomes. As we depart from these conditions, we see significant differences in the outcomes of the three methods. The BV outcomes are most sensitive to the amplitude and the frequency of the perturbations. Regardless of the amplitude of the perturbations, the EBV outcomes are structurally unambiguous and robust.

We revisit the CY92 simulations of BV and EBV, using the same forcing and initial conditions used Section 4.2 but set the time scale parameter $\alpha = 0.01$, we increase the spatial resolution to 320 points, and set the integration time step at $\delta t = 0.000015$. The perturbation field is the same as in (45); however, we will increase the size of the perturbation of each of the components. For $\epsilon_j = 0.6, 0.8$, and 1.2 , with $j = 1, 2, \dots, 6$, in (45) we obtain Figure 9, which shows all BVs at $t = 19.025$. We expect a simple structure in the perturbation field and thus the BV should reflect this. The EBV results appear in Figure 10. Comparison of Figures 9 and 10 show how the BV outcomes look qualitatively different from their EBV counterparts as we increase the initial amplitude of the ensemble. The comparison is further aided by reference to Figure 11, in which each ensemble member of the EBV, at $t = 19.025$, has been rescaled to 1 in amplitude. The BV outcomes shown in Figure 9 do not yield the structural clarity that the EBV ensemble displays in Figure 10; this is a natural consequence of the size-ordering inherent in the EBV algorithm.

In reference to Figure 9c and Figure 10c we see the similarity between one and only one of the BV vectors and the largest EBV. A striking structural feature in the EBV results shown in Figure 10 is that there are only 3 significantly-different shapes, among all of the six vectors (see Figure 11, where the vector have been rescaled to have size 1 in L^2 .) For relative comparison, see Figure 9.

The sensitivity of the outcomes to the size of initial perturbations is significantly different in the EBV and the BV outcomes. There are shape variations among the BVs as the amplitude of the perturbation is increased,

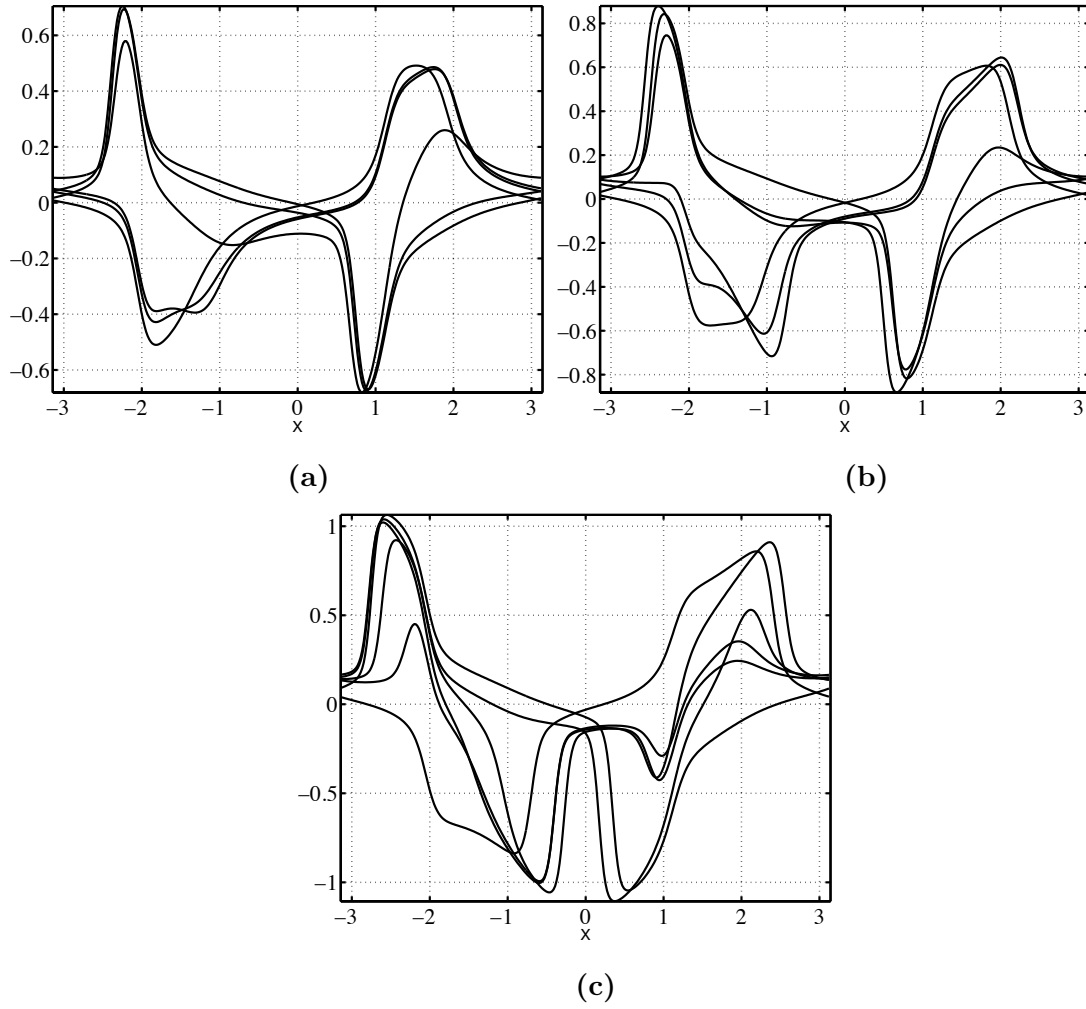


Figure 9: *BVs at $t = 19.025$, corresponding to $\epsilon_j = \epsilon$, for all j . (a) $\epsilon = 0.6$, (b) $\epsilon = 0.8$, (c) $\epsilon = 1.2$. The outcomes are very sensitive to nonlinearity. There is a resulting ambiguity in structure of the perturbation field.*

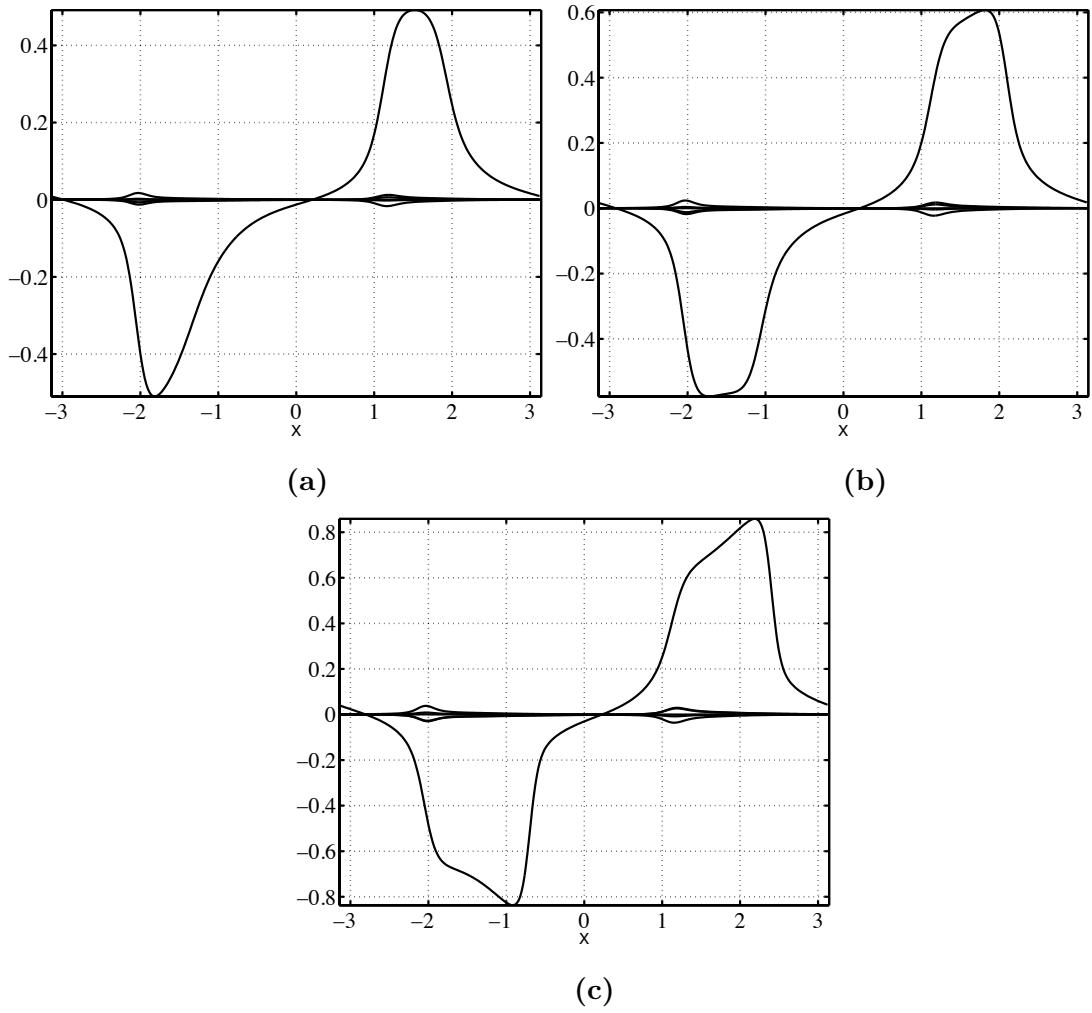


Figure 10: *EBV* with (a) $\epsilon = 0.6$, (b) $\epsilon = 0.8$, (c) $\epsilon = 1.2$. Compare outcomes to Figure 9. The vectors are shown in their original scales.

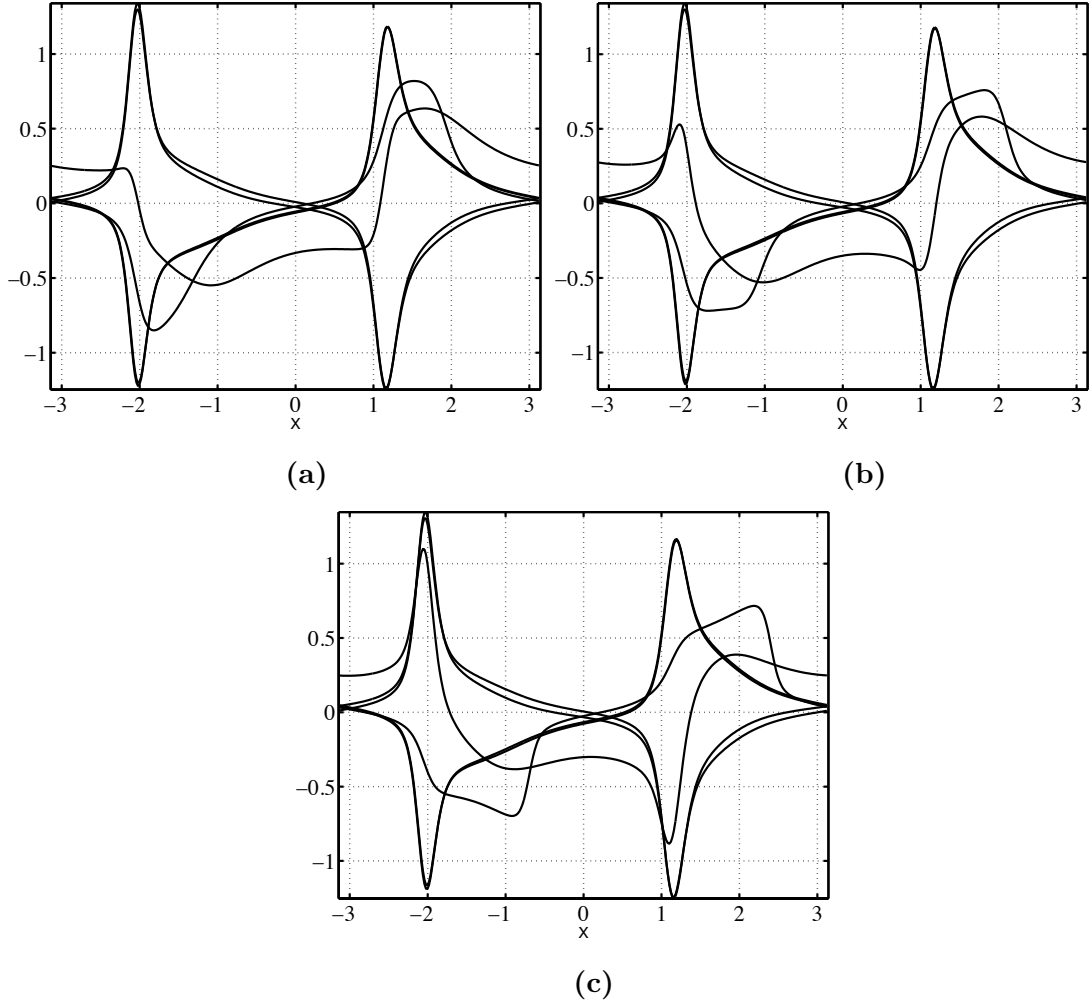


Figure 11: *EBV outcomes shown in Figure 10, with each vector rescaled to an L^2 -norm of 1 to aid in visual comparison. Perturbation amplitudes (a) $\epsilon = 0.6$, (b) $\epsilon = 0.8$, and (c) $\epsilon = 1.2$. Note that as the perturbation amplitude increases, the resemblance between the BV and the EBV outcomes is lost, except for one of the vectors.*

as evidenced in Figure 9. This degree of sensitivity is not as prevalent in the EBV outcomes, shown in Figure 11. The BV algorithm does not distinguish between different perturbation scales. A large perturbation BV calculation will thus not yield information concerning smaller scales. In contrast, see Figure 11c, for the EBV case, where small scale information is evident. (See also Figure 10c).

5 Implementation Issues

Like the BV algorithm, the EBV algorithm is capable of dealing with legacy code. The important difference between implementing a code that does an ensemble of BV and the EBV is that while the former can be run concurrently, in the EBV the ensemble members require normalization to each other. In terms of coding, this is a minor issue, if the state variable dimension of the underlying model is moderate. For very large problems communication becomes an issue, but not at all unfamiliar in concurrent or hybrid computing. Any additional computational issues borne by the EBV are well outweighed by the higher informational content of the EBV over the BV. Moreover, the EBV is much more robust under the nonlinear effects, as illustrated in Section 4, than BV. This last aspect is very important and it should be studied in the context of more chaotic equations in fluid mechanics, meteorology, and geophysics.

The numerical outcome of the BV and EBV algorithms depends on the choice of norm used for rescaling. This is alluded-to in Rivière et al. (2008), but the reason for this dependence turns out to be easily explained and can be significant if nonlinear effects are not negligible. The dependence of the outcomes on the norm lies in the fact that it is not possible, in the general case, to scale out the norm in the algorithm if the underlying dynamics are nonlinear. To illustrate the norm dependence, we consider a simple 2-dimensional system

$$\begin{aligned}\frac{dX_1}{dt} &= X_2, \\ \frac{dX_2}{dt} &= -\sin(8X_1), \quad t > 0,\end{aligned}\tag{57}$$

subject to the (same) initial conditions $(X_1(0), X_2(0)) = (0.8, -1)$. Figure 12 shows the outcomes of the BV algorithm for three different choices of norms, starting from perturbations of size no greater than 0.15. The norms we employ are all standard (finite) L_p norms. The resulting BVs reflect the characters of these norms. That different outcomes are obtained is not surprising: the shape of the unit sphere changes depending on which L_p norm is used.

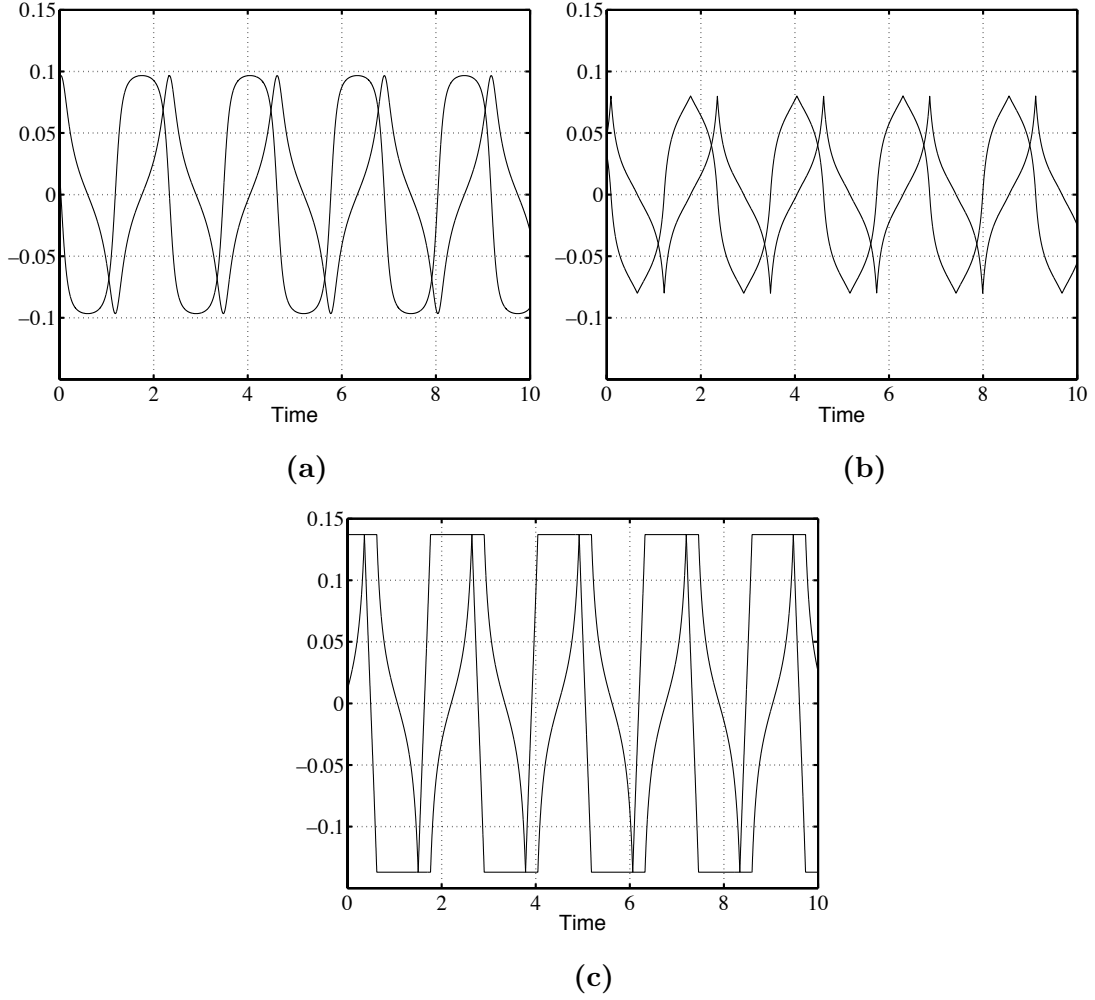


Figure 12: (a) L_2 -norm, (b) L_1 -norm, and (c) L_∞ -norm of a sample BV as a function of time, for the system in (57). The initial conditions are the same in all cases, the amplitude of the perturbations was 0.15.

A more weighty consideration related to norms concerns physical and theoretical considerations. Conservation laws provide guidance for the most appropriate norms for the given dynamics (for instance, in Rayleigh-Bénard convection, the temperature enjoys a maximum principle, while the velocity does not), but other considerations will play a role (*e.g.*, a sup-norm may be an obvious choice in determining the location of severe weather events). It will not be uncommon, thus, that a mixture of norms may be necessary in multi-physics problems; the fact that the choice of norm in the EBV/BV affects the results is in fact a good thing, not a bad one.

To end this section we wish to highlight a practical consideration that may be not be familiar to practitioners, implementing BV or EBV algorithms computationally. The specific issue is the impact of finite precision computing on the outcomes. It is easy to show that these algorithms have high numerical sensitivity. In order to illustrate how this plays out we will consider the problem of calculating the vectors associated with

$$\frac{dX}{dt} = AX, \quad t > 0, \quad X(0) = X_0. \quad (58)$$

Here A is a square matrix of constants. For linear problems BV, EBV (and the finite-time Lyapunov) algorithms must yield the same vectors. We will choose to illustrate computationally numerical ill-conditioning on a problem that exhibits transient growth due to the non-normal structure of the matrix A . We emphasize, however, that the numerical sensitivity we will be highlighting in this example does not hinge on the nature of the dynamics, but rather, on the algorithmic form of the BV and EBV themselves.

We take A to be an upper-triangular Jordan-block matrix of dimension 5, where A has a single eigenvalue λ , which is repeated on the main diagonal, while the diagonal directly above the main diagonal has non-zero entries. For simplicity we assume these to have the same value ρ . We assume further that $\lambda = -1$ and $\rho = 1$. The general solution operator contains “generalized” eigen-solutions with growth rates $t^k e^{-t}$, where $0 \leq k \leq 4$. (In this case, the eigenvalue has (algebraic) multiplicity 5.) For this linear problem we expect to see, provided we take t sufficiently long to forget the transient, a very trivial outcome to BV, or EBV.

In Figure 13 we summarize the results of the BV calculation on this system. We employed an integration time step of 0.001, a perturbation initially of magnitude 0.01. We performed the calculation in double precision, using an explicit Runge-Kutta 4, but paid little attention to how numerical sensitivity was handled. In Figure 13a we show the BV, *i.e.*, \mathcal{Y} , as a function of time. The calculation of BV, using (3), clearly diverged, shortly after about $t = 30$ (and thus not shown in Figure 13a). However, there were indications that something was not right even before it became obvious that the solution

was wrong: In Figure 13b we show the 2-norm of $M(Y_n + \delta\mathcal{Y}_n, \delta t) - M(Y_n, \delta t)$. It monotonically increases, even though all the factors $t^k e^{-t}$, for $0 \leq k \leq 4$ in the solution operator, decay for $t > 4$.

The outcome is not related to choosing a rescaling time to be too long: in fact, in this computation the rescaling is performed at each computational step. Moreover, the time step chosen was sufficiently small to guarantee asymptotic stability in the numerical integrator. A smaller time step would have been able to ameliorate to a certain extent the numerical sensitivity of the difference $M(Y_n + \delta\mathcal{Y}_n, \delta t) - M(Y_n, \delta t)$, however, in very large scale problems this might not be practical or even possible. Figure 13c shows the 2-norm of $M(Y_n + \delta\mathcal{Y}_n, \delta t) - M(Y_n, \delta t)$, using a numerically well-conditioned implementation of the BV algorithm. The strategy used to obtain a well-conditioned outcome was to normalize the elements in the required subtraction before computing their difference. The well-conditioned calculation was capable of qualitatively good results. However, the strategy adapted here to increase the numerical stability was by no means generally applicable to all problems, nor was it optimal.

6 Concluding Remarks

The main thrust of our work is to propose an ensemble-based vector breeding algorithm, the Ensemble Bred Vector (EBV) algorithm. It is based on the Bred Vector (BV) algorithm introduced by Toth and Kalnay (1993, 1997). We compare the EBV to the BV algorithm and the finite-time Lyapunov Vector algorithms. In the EBV, an ensemble of initial perturbations is bred concurrently and then rescaled by the size of the largest member of the bred ensemble. The uniform normalization of all the ensemble members after each cycle is the distinctive trait of the EBV algorithm, which leads to some profound differences when compared to the BV algorithm.

As expected, when initial perturbations are sufficiently small, the EBV, the BV, and the finite-time Lyapunov vector algorithms lead to similar results. We gave a theoretical justification of this phenomenon by looking at the corresponding time-continuum analogues of the BV and EBV algorithms. We rescale frequently: The algorithms and results are formulated, for simplicity, assuming that the rescaling is done after every discrete time step.

In Section 3, we develop a solid mathematical basis for both the BV and EBV algorithms and show that each algorithm results in good approximations of the solutions of the tangent linear model, when the step-size δt is small. As is seen in Table 1, the EBV algorithm has a substantial advantage over the (classical) BV algorithm, in the sense that the drop in the minimal EBV error is substantially better than the corresponding drop in the minimal

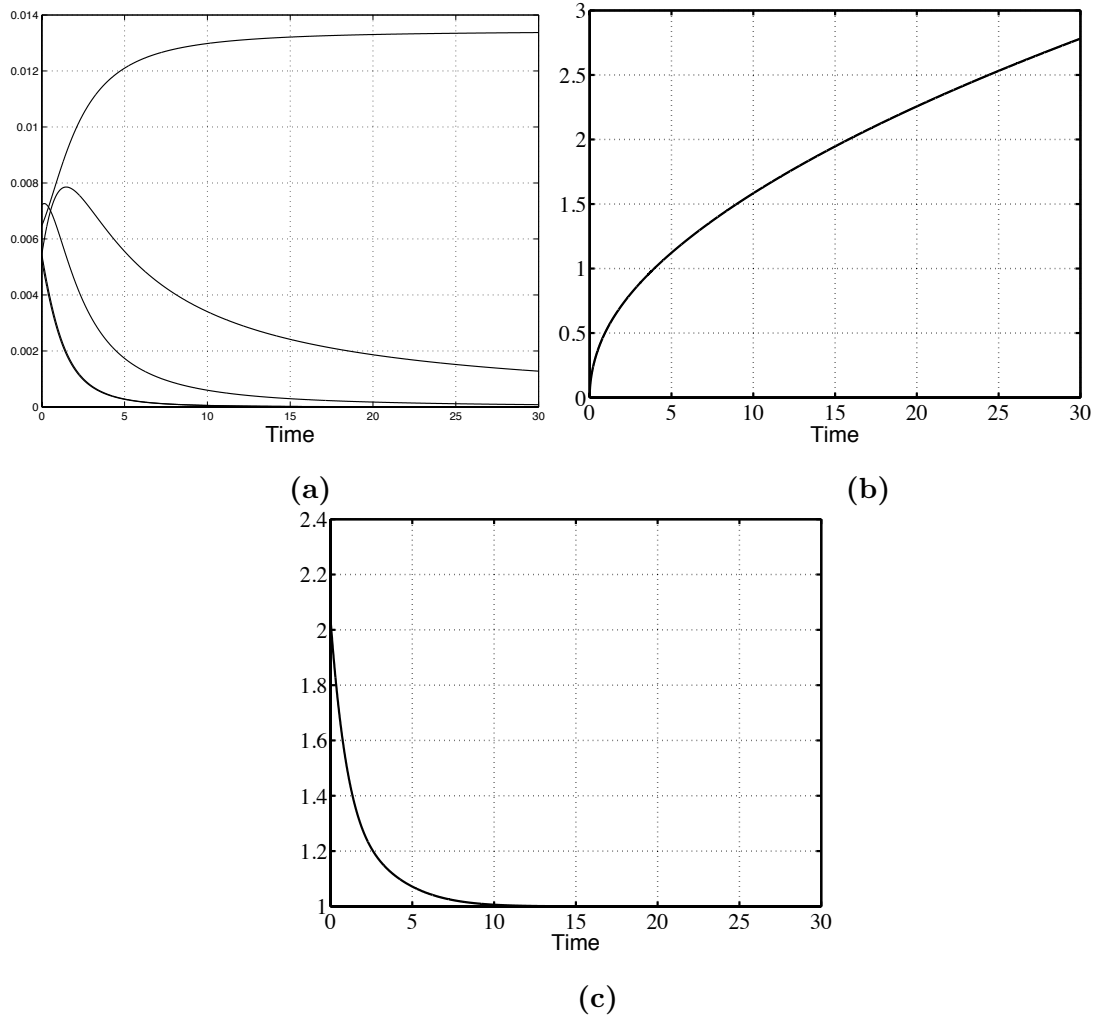


Figure 13: (a) A sample BV, as a function of time. The calculation diverges and eventually fails (not shown); (b) 2-norm of the numerically-approximated $M(Y_n + \delta\mathcal{Y}_n, \delta t) - M(Y_n, \delta t)$ as a function of time, for the linear system $\frac{dX}{dt} = AX$. Here, A is an upper-triangular Jordan-block matrix of dimension 5, where A has a single eigenvalue -1 , which is repeated on the main diagonal, while the superdiagonal has entries 1. (c) 2-norm of the numerically-approximated $M(Y_n + \delta\mathcal{Y}_n, \delta t) - M(Y_n, \delta t)$, well-conditioned case.

BV error.

In the study of the Lorenz attractor, the classical BV algorithm has a shortcoming which limits one's ability to use this algorithm to study the dynamics inside the attractor. In particular, equation (5) implies that $\|\delta\mathcal{Y}_{n+1}(\iota)\| = \epsilon$, for all $n \geq 0$ and all $\iota \in \mathbb{I}$. The BV algorithm maps the cloud onto a 2-dimensional sphere. The third dimension is lost, and with it so are the fractal patterns seen in Figures 4 and 1. Also the many patterns seen in Figure 2 are lost because the BV alternative would map everything onto the single line at height ϵ .

Lastly, in Figures 9, 10, and 11, we examine a series of related test problems that depart from the tangent linear model. The point here is to get a comparison of the performance of the two algorithms, as one moves further into nonlinear regime. Once again, one sees a significant advantage of the EBV over the BV.

The theoretical aspects for the EBV and BV algorithms are presented in Subsection 2.5. As is noted there, our application to the basic issue of the sensitivity with respect to errors in the initial conditions relies heavily on the Johnson, et al manuscript, JPS87.

In conclusion, for the applications described in this article, we believe that the new EBV algorithm has been shown to be superior to the traditional BV algorithm. Finally, we ask: Is the EBV algorithm the "last" word on modifications of the classical BV? Probably not. However, we do expect that the EBV will serve as a good starting point for new theories of bred vectors.

Acknowledgements

JMR was partially supported by NSF grant DMS-0335360. ALM was partially supported by NSF grant DMS-0708902, DMS-1009713 and DMS-1009714. NB, ALM, JMR, and GRS wish to thank the Institute for Mathematics and its Applications (IMA) for their support and hospitality. Research at the IMA is supported by the National Science Foundation and the University of Minnesota. GRS acknowledges his appreciation of Robert Sacker for related suggestions he made at an early stage in the development of this article. We express our sincere gratitude to the referees, who made an in depth reading of the original manuscript and offered several helpful suggestions which led to improvements in the paper.

References

- L. Arnold. *Random Dynamical Systems*. Springer, New York, 1998.
- R. Buizza, J. Tribbia, F. Molteni, and T. Palmer. Computation of optimal unstable structures for numerical weather prediction models. *Tellus*, 45A: 388–407, 1993.
- P. Cessi and W. R. Young. Multiple equilibria in two-dimensional thermohaline circulation. *Journal of Fluid Mechanics*, 241:291–309, 1992.
- K. K. W. Cheung. Ensemble forecasting of tropical cyclone motion: comparison between regional bred modes and random perturbations. *Meteorology and Atmospheric Physics*, 78:23–35, 2001.
- M. Corazza, E. Kalnay, D. J. Patil, S.-C. Yang, R. Morss, M. Cai, I. Szunyogh, B. R. Hunt, and J. A. Yorke. Use of the breeding technique to estimate the structure of the analysis errors of the day. *Nonlinear Processes in Geophysics*, 10:233–243, 2003.
- B. Deremble, F. DAndrea, and M. Ghil. Fixed points, stable manifolds, weather regimes, and their predictability. *Chaos*, 19:043109, 2009.
- E. Evans, N. Bhatti, J. Kinney, L. Oann, M. Peña, S. Yang, and E. Kalnay. RISE undergraduates find that regime changes in Lorenz model are predictable. *Bulletin of the American Meteorological Society*, 85:520–524, 2004.
- G. Eyink. Statistical hydrodynamics of the thermohaline circulation in a two dimensional model. *Tellus, A*, 57:100–115, 2005.
- C. Foias, G. R. Sell, and R. Temam. Inertial manifolds for nonlinear evolutionary equations. *Journal of Differential Equations*, 73:309 – 353, 1988.
- T. Gneiting and A. E. Raftery. Atmospheric science - weather forecasting with ensemble methods. *Science*, 310:248–249, 2005.
- S. Hallerberg, D. Pazo, J. M. Lopez, and M. A. Rodriguez. Logarithmic bred vectors in spatiotemporal chaos: Structure and growth. *Physical Review E*, 81, 2010.
- J. A. Hansen and L. A. Smith. The role of operational constraints in selecting supplementary observations. *Journal of the Atmospheric Sciences*, 57: 2859–2971, 2000.

- R. A. Johnson, K. J. Palmer, and G. R. Sell. Ergodic properties of linear dynamical systems. *SIAM Journal of Mathematical Analysis*, 18:1–33, 1987.
- E. Kalnay. *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, Cambridge, 2003.
- J. P. LaSalle and S. Lefschetz. *Stability by Lyapunov’s Direct Method with Applications*. Academic Press, New York, 1961.
- E. N. Lorenz. Deterministic nonperiodic flow. *Journal of Atmospheric Science*, 20:130–141, 1963.
- A. M. Lyapunov. The general problem of the stability of motion. *International Journal of Control*, 55(3):521–790, 1992. Translated by A. T. Fuller from Édouard Davaux’s French translation (1907) of the 1892 Russian original, (historical introduction) by Fuller, a Smirnov, and the bibliography Barrett,
- B. B. Mandelbrot. *Fractals: form, chance, and dimension*. W. H. Freeman and Co., San Francisco, Calif., revised edition, 1977. Translated from the French.
- M. Mu and Z. N. Jiang. A new approach to the generation of initial perturbations for ensemble prediction: Conditional nonlinear optimal perturbation. *Chinese Science Bulletin*, 53:2062–2068, 2008.
- T. N. Palmer, R. Gelaro, J. Barkmeijer, and R. Buizza. Singular vectors, metrics and adaptive observations. *Journal of Atmospheric Sciences*, 55: 633–653, 1998.
- V. A. Pliss and G. R. Sell. Robustness of exponential dichotomies in infinite-dimensional dynamical systems. *Journal of Dynamics and Differential Equations*, 11:471–513, 1999.
- C. Primo, M. A. Rodriguez, and J. M. Gutierrez. Logarithmic bred vectors. a new ensemble method with adjustable spread and calibration time. *Journal of Geophysical Research-Atmospheres*, page D05116, 2008.
- O. Rivière, G. Lapeyre, and O. Talagrand. Nonlinear generalization of singular vectors: Behavior in a baroclinic unstable flow. *Journal of the Atmospheric Sciences*, 65:1896–1911, 2008.
- R. J. Sacker and G. R. Sell. Existence of dichotomies and invariant splittings for linear differential systems. II. *Journal of Differential Equations*, 22(2): 478–496, 1976.

- R. J. Sacker and G. R. Sell. Lifting properties in skew-product flows with applications to differential equations. *Memoirs of the American Mathematical Society*, 11(190):iv+67, 1977.
- R. J. Sacker and G. R. Sell. A spectral theory for linear differential systems. *Journal of Differential Equations*, 27(3):320 – 358, 1978.
- R. J. Sacker and G. R. Sell. The spectrum of an invariant submanifold. *Journal of Differential Equations*, 38(2):135 – 160, 1980.
- G. R. Sell and Y. You. *Dynamics of Evolutionary Equations*. Springer, New York, 2002.
- Z. Toth and E. Kalnay. Ensemble forecasting at NCEP: the generation of perturbations. *Bulletin of the American Meteorological Society*, 74:2317–2330, 1993.
- Z. Toth and E. Kalnay. Ensemble forecasting at NCEP: the breeding method. *Monthly Weather Review*, 125:3297–3318, 1997.
- X. G. Wang and C. H. Bishop. A comparison of breeding and ensemble transform Kalman filter ensemble forecast schemes. *Journal of the Atmospheric Sciences*, 60:1140–1158, 2003.
- M. Z. Wei and Z. Toth. A new measure of ensemble performance: Perturbation versus error correlation analysis (PECA). *Monthly weather Review*, 131:1549–1565, 2003.
- C. L. Wolfe and R. M. Samelson. An efficient method for recovering Lyapunov vectors from singular vectors. *Tellus*, 59A:355–366, 2007.